

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 14:42:49 ; Search time 9009 Seconds
(without alignments)
11036.752 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcctacagcaca.....gcataacattctaaatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hlg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2052	100.0	2052	6 E40087	E40087 Plant promo
2	2048.8	99.8	2052	6 E40093	E40093 Plant promo
3	2038	99.3	2056	6 E40091	E40091 Plant promo
4	2034	99.1	2048	6 E40090	E40090 Plant promo
5	2032.4	99.0	2048	6 E40089	E40089 Plant promo
6	241.2	11.8	2042	6 AR076817	AR076817 Sequence
7	241.2	11.8	2042	6 E15125	E15125 Promoter. 7
8	196.6	9.6	247	6 AR076816	AR076816 Sequence
9	196.6	9.6	247	6 E15124	E15124 Promoter. 7
10	195.6	9.5	246	6 AR146852	AR146852 Sequence
11	195.6	9.5	246	6 E55065	E55065 Plant promo
12	174	8.5	8005	8 DCAL8706	Y18706 Daucus caro
13	162.4	7.9	2831	6 BD188672	BD188672 Promotor
14	162.4	7.9	2865	6 BD188674	BD188674 Promotor
15	162.4	7.9	4886	8 DARGCHS2	D1625 Carrot qCHS
16	119	5.8	140	6 AR146856	AR146856 Sequence
17	119	5.8	140	6 AR146857	AR146857 Sequence
18	119	5.8	140	6 E55069	E55069 plant promo
19	119	5.8	140	6 E55070	E55070 plant promo

C 20	106.6	5.2	250029	3	AE014839	Plasmodi
C 21	103.6	5.0	172816	9	AC093899	Homo sapi
C 22	101.8	5.0	258658	3	AE014832	Plasmodi
C 23	100.4	4.9	348174	3	CR382399	Plasmodi
C 24	99.4	4.8	8056	6	AX599046	Sequence
C 25	99.2	4.8	205130	2	AC105425	Homo sapi
C 26	97	4.7	14867	3	AE001398	Plasmodi
C 27	96.6	4.7	67970	3	PFMAL1P3	Plasmodi
C 28	96.4	4.7	4601	3	DMU11584	U1584 Drosophila
C 29	96.4	4.7	19517	3	DMU37541	U37541 Drosophila
C 30	95.8	4.7	64394	9	AL928596	Human DNA
C 31	95.2	4.6	313050	3	PFAL929352	Plasmodi
C 32	93.8	4.6	143331	9	AC091214	Homo sapi
C 33	93	4.5	349980	6	AX344555	Sequence
C 34	92.6	4.5	180629	2	CR388025	Danio rer
C 35	92.6	4.5	349751	3	PFMAL4P3	Plasmodi
C 36	92.4	4.5	261771	2	CR356223	Danio rer
C 37	91.8	4.5	8056	6	AX598900	Sequence
C 38	91.6	4.5	8056	6	AX599046	Sequence
C 39	91	4.4	21153	8	YSCWTCG13	L36897 Saccharomyc
C 40	91	4.4	118642	9	AC126283	Homo sapi
C 41	91	4.4	178427	2	CR392006	Danio rer
C 42	90.8	4.4	86826	3	PFMAL3P5	Plasmodi
C 43	90.6	4.4	155106	9	AC104069	Homo sapi
C 44	90.6	4.4	178670	9	AC104073	Homo sapi
C 45	90.6	4.4	206038	2	CR847802	Danio rer

ALIGNMENTS

RESULT 1	E40087	2052 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	E40087	Plant promoter and terminator.			
DEFINITION	E40087				
ACCESSION	E40087.1	GI:18627203			
VERSION	JP 2000166577-A/1.				
KEYWORDS	unidentified				
SOURCE	unclassified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 2052)				
AUTHORS	Nishikawa,S. and Oeda,K.				
TITLE	Plant promoter and terminator				
JOURNAL	Patent: JP 2000166577-A 1 20-JUN-2000;				
COMMENT	SUMITOMO CHEM CO LTD				
	OS Daucus carota L.				
	PN JP 2000166577-A/1				
	PD 20-JUN-2000				
	PF 01-OCT-1999				
	PR JP 1999281475				
	PI SATOMI NISHIKAWA, KENJI OEDA				
	PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1.91), PC				
	C12N15/00,				
	PC C12N5/00, (C12N5/00,C12R1.91)				
	CC Key Location/Qualifiers				
	FT promoter (1)..(2052).				
	source Location/Qualifiers				
	1..2052				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				

Query Match 100.0%; Score 2052; DB 6; Length 2052;
Best Local Similarity 100.0%; Pred. No. 4.4e-310;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGTGTGCCTACACACATAGGCTGTTGGTTGAGAGACACAGCTGCTCTGA 60
|||||
Db 1 CATGTGTGCCTACACACATAGGCTGTTGGTTGAGAGACACAGCTGCTCTGA 60

Qy	61	CTTCTCTCTCTCTTTGACCTGTTGTTATAAAGAGTAGAATAATTTTTAAAAAGCTCGAA	120
Db	61	CTTCTCTCTCTTTGACCTGTTGTTATAAAGAGTAGAATAATTTTTAAAAAGCTCGAA	120
Qy	121	TACTAACTCTCTCTCACAACTTCCGCTCTTTTCCAAACACCTTATTAACCTTTTTTACT	180
Db	121	TACTAACTCTCTCTCACAACTTCCGCTCTTTTCCAAACACCTTATTAACCTTTTTTACT	180
Qy	181	TCTCAATTTCTACCTCCACTCTTTGCTATAGCAAGAAATCACCTCTTTTAAAGCTAACCCA	240
Db	181	TCTCAATTTCTACCTCCACTCTTTGCTATAGCAAGAAATCACCTCTTTTAAAGCTAACCCA	240
Qy	241	AACGGCTCAATAAAGAGCAATTCATAAATGTATCTTTCAAATTTTAGGATAACAATACGT	300
Db	241	AACGGCTCAATAAAGAGCAATTCATAAATGTATCTTTCAAATTTTAGGATAACAATACGT	300
Qy	301	GAACAGGGTATTTTTTAAACGGTCAACAAATTCATAAATCTAATAATTTTAAAGCTGGAACA	360
Db	301	GAACAGGGTATTTTTTAAACGGTCAACAAATTCATAAATCTAATAATTTTAAAGCTGGAACA	360
Qy	361	CCGCTCTCCAGGATAATATATTTTAAATTTGTAGCCTCCCTTTTAAACCAATTCGCATGC	420
Db	361	CCGCTCTCCAGGATAATATATTTTAAATTTGTAGCCTCCCTTTTAAACCAATTCGCATGC	420
Qy	421	AGGACGACTTAGGTGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGAACAGTGGTTC	480
Db	421	AGGACGACTTAGGTGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGAACAGTGGTTC	480
Qy	481	ATGCTCAGCCATCAAAATTTGACAAAAACCGACACAACACTCTATCCAGCTACTATCTTT	540
Db	481	ATGCTCAGCCATCAAAATTTGACAAAAACCGACACAACACTCTATCCAGCTACTATCTTT	540
Qy	541	TGCGCGAATGCTCTCAAAATGTTTTTATATGTAATAATGCTTAAATGCTTAAAGGATAAGT	600
Db	541	TGCGCGAATGCTCTCAAAATGTTTTTATATGTAATAATGCTTAAATGCTTAAAGGATAAGT	600
Qy	601	AAATTTCCCGTTTAAACAGTTTCTTAATATATATGTTTACACTTACAAAGAGATATTCGT	660
Db	601	AAATTTCCCGTTTAAACAGTTTCTTAATATATATGTTTACACTTACAAAGAGATATTCGT	660
Qy	661	AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT	720
Db	661	AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT	720
Qy	721	TGCTACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTA	780
Db	721	TGCTACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTA	780
Qy	781	ATTAGAGCTAATTAATTAAGTTACTAATAAATAAGAGAGGTAGTAAACAGAAAGCAGGTA	840
Db	781	ATTAGAGCTAATTAATTAAGTTACTAATAAATAAGAGAGGTAGTAAACAGAAAGCAGGTA	840
Qy	841	AAAAAAGAGCTTGCTGCTGTTGTTAGTTGTTGAGCTCATTTCTTTTAAAGTAATG	900
Db	841	AAAAAAGAGCTTGCTGCTGTTGTTAGTTGTTGAGCTCATTTCTTTTAAAGTAATG	900
Qy	901	TAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTCAAGAAATTTATA	960
Db	901	TAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTCAAGAAATTTATA	960
Qy	961	TTAAACGAAATCATTTTATAACATGCTCTCCGCTGTCATTATATAGGATCACTTAC	1020
Db	961	TTAAACGAAATCATTTTATAACATGCTCTCCGCTGTCATTATATAGGATCACTTAC	1020
Qy	1021	TGATCATCCATTAACCTTTGTTTAAACAAATTCATAGATAAAAATATCTTCAATGAA	1080
Db	1021	TGATCATCCATTAACCTTTGTTTAAACAAATTCATAGATAAAAATATCTTCAATGAA	1080
Qy	1081	AAGAAGCAATGTCTCTTTTGAAGAAAAAATAAGTACTCCCTCCGCTCTCTGAAATGT	1140
Db	1081	AAGAAGCAATGTCTCTTTTGAAGAAAAAATAAGTACTCCCTCCGCTCTCTGAAATGT	1140
Qy	1141	ATACATATGATTTGGACACGGAGCTAAGAAAAATGTATAAAGTAAATGTAGATAAAG	1200

Db	1141	ATACATATGATTTGGACACGGAGACTTAAGAAAAATGTAATAAGTAAATGTAGAGTAAAAAG	1200
Qy	1201	AAAGGAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAAATGTAGATTTAG	1260
Db	1201	AAAGGAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAAATGTAGATTTAG	1260
Qy	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATAAAAAATTTACTATTTTG	1320
Db	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATAAAAAATTTACTATTTTG	1320
Qy	1321	AGAAAGTTTGGAAATGTATAGAAATTTAGTGGGACATCCATAAAGAAAGTGTATAGAAT	1380
Db	1321	AGAAAGTTTGGAAATGTATAGAAATTTAGTGGGACATCCATAAAGAAAGTGTATAGAAT	1380
Qy	1381	TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTGTTATTTGATTCTCA	1440
Db	1381	TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTGTTATTTGATTCTCA	1440
Qy	1441	TAAAGTATATAAATCTATGTTATATAATGATAATAAATTTTAAAAAATACTATATTAAT	1500
Db	1441	TAAAGTATATAAATCTATGTTATATAATGATAATAAATTTTAAAAAATACTATATTAAT	1500
Qy	1501	CTGATTAGTCGATTACCGCTTTTATAATTTTACAATACTGAGTAAATGAAATAAATCAG	1560
Db	1501	CTGATTAGTCGATTACCGCTTTTATAATTTTACAATACTGAGTAAATGAAATAAATCAG	1560
Qy	1561	TTATCTGAAAAAGCAATAATATCTTTGTAAAAACAGCGTTCCGTCAAAATGGGAAGTTCATG	1620
Db	1561	TTATCTGAAAAAGCAATAATATCTTTGTAAAAACAGCGTTCCGTCAAAATGGGAAGTTCATG	1620
Qy	1621	TGTTTCAATAGCTTTTAAATATAAAGTAAATTTTAAATTAATTTTGTGTTTTCAGA	1680
Db	1621	TGTTTCAATAGCTTTTAAATATAAAGTAAATTTTAAATTAATTTTGTGTTTTCAGA	1680
Qy	1681	AATTTAAAAATAAATTAATGAGCATGGGAAGTTTCACGGGCATCATTTGAGCAGCACTAGACT	1740
Db	1681	AATTTAAAAATAAATTAATGAGCATGGGAAGTTTCACGGGCATCATTTGAGCAGCACTAGACT	1740
Qy	1741	GTTTGAACAATGTATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATAAT	1800
Db	1741	GTTTGAACAATGTATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATAAT	1800
Qy	1801	GCATTTCAATAACATCTTTTCAAAATTTTCAAAACACAGCTTTTAACTTTTCTTCAACG	1860
Db	1801	GCATTTCAATAACATCTTTTCAAAATTTTCAAAACACAGCTTTTAACTTTTCTTCAACG	1860
Qy	1861	GATTGGAATCCTTTTCTAAACCTTTTAAATAAAAAATGCAATTAATGTAATTTTATC	1920
Db	1861	GATTGGAATCCTTTTCTAAACCTTTTAAATAAAAAATGCAATTAATGTAATTTTATC	1920
Qy	1921	AACACCTCAACATTTAGTTAGGTAATTAATAGGTGCTCTTGGTGTCTACTATCAT	1980
Db	1921	AACACCTCAACATTTAGTTAGGTAATTAATAGGTGCTCTTGGTGTCTACTATCAT	1980
Qy	1981	CACATCAATCTTACACCAAAAACCTTGAGCTAAATTTTCTACTTATTTCTCAGCAATAAC	2040
Db	1981	CACATCAATCTTACACCAAAAACCTTGAGCTAAATTTTCTACTTATTTCTCAGCAATAAC	2040
Qy	2041	ATTCTAAATATC 2052	
Db	2041	ATTCTAAATATC 2052	

RESULT 2
E40093
LOCUS E40093 2052 bp DNA linear PAT 31-JAN-2002
DEFINITION Planc promoter and terminator.
ACCESSION E40093
VERSION E40093.1 Gr:18627209
KEYWORDS JP 2000166577-A/7.
SOURCE unidentified
ORGANISM unidentified

REFERENCE AUTHORS TITLE JOURNAL COMMENT	unclassified. 1 (bases 1 to 2052) Nishikawa, S. and Oeda, K. Plant promoter and terminator Patent: JP 2000166577-A 7 20-JUN-2000; SUMITOMO CHEM CO LTD OS Daucus carota L. PN JP 2000166577-A/7 PD 20-JUN-2000 PF 01-OCT-1999 JP 1999281475 PR PI SATOMI NISHIKAWA, KENJI OEDA PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//C12N5/10,C12R1/91, PC C12N15/00, PC C12N5/00, (C12N5/00,C12R1:91) CC FH FT	Key promoter	Location/Qualifiers (1). (2052).	Location/Qualifiers (1). (2052).	Query Match Best Local Similarity 99.8%; Score 2048.8; DB 6; Length 2052; Matches 2050; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
ORIGIN	1 CATGTGCGCCACACACATAGGGCGCTGTTGGTTGAGAGACGAGAAAGCTGCTCTGA 60 1 CATGTGCGCCACACACATAGGGCGCTGTTGGTTGAGAGACGAGAAAGCTGCTCTGA 60 61 CTTCTCTCTCTTTGACCTGTTGTATAAAGAAAGTAGAAATATTTTAAAGAGCTGCGAA 120 61 CTTCTCTCTCTTTGACCTGTTGTATAAAGAAAGTAGAAATATTTTAAAGAGCTGCGAA 120 121 TACTAACTTCTCTCTCACAACTTCGCGCTCTTTTCCAAACACCTTTATTAACCTTTTACT 180 121 TACTAACTTCTCTCTCACAACTTCGCGCTCTTTTCCAAACACCTTTATTAACCTTTTACT 180 181 TCTCATTTCT 240 181 TCTCATTTCT 240 241 AAGCGCTCAATAAAGATCATTCAATAATGATCTTTCAATTTTATAGATTAACCAATAGT 300 241 AAGCGCTCAATAAAGATCATTCAATAATGATCTTTCAATTTTATAGATTAACCAATAGT 300 301 GAACAGGGTATTTTAAACGCTGCAACAAATTTCTAATAATTTTACCTGGCGGCTGAACA 360 301 GAACAGGGTATTTTAAACGCTGCAACAAATTTCTAATAATTTTACCTGGCGGCTGAACA 360 361 CGCTCTCCAAGATAATATATTTTAAATTTTATAGTCTCTCTCTCTCTCTCTCTCTCTCT 420 361 CGCTCTCCAAGATAATATATTTTAAATTTTATAGTCTCTCTCTCTCTCTCTCTCTCTCT 420 421 AGGACGACTTAGTGGAATACATTTGTAAGTCTTTTAAACAAAGAAAGAGTGGTTC 480 421 AGGACGACTTAGTGGAATACATTTGTAAGTCTTTTAAACAAAGAAAGAGTGGTTC 480 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACCACTCTATCCAGTACTATCTTT 540 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACCACTCTATCCAGTACTATCTTT 540 541 TGCGCGAATGCTCTCTCAAAATGTTTTTATATGTAATAATATGTCCTCAAGGATAGT 600 541 TGCGCGAATGCTCTCTCAAAATGTTTTTATATGTAATAATATGTCCTCAAGGATAGT 600 601 AAAATTCGCGTTTAAACGAGTTGTTTAAATATATGTTTACACTTCAAGAGATATTCGT 660 601 AAAATTCGCGTTTAAACGAGTTGTTTAAATATATGTTTACACTTCAAGAGATATTCGT 660				

Db	1741	GTTGAACAATGATGTCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT	1800
Qy	1801	GCATTTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACG	1860
Db	1801	GCATTTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACG	1860
Qy	1861	GATTGGAATCCTTTTCTAAACTTTTAAATAAATAAATAATGCAATATTGTAATATTATC	1920
Db	1861	GATTGGAATCCTTTTCTAAACTTTTAAATAAATAAATAATGCAATATTGTAATATTATC	1920
Qy	1921	AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT	1980
Db	1921	AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT	1980
Qy	1981	CACATCAATCTTACACCAAACTTGGAGCTTAATTTTCTACTATTCTCAGCAATAAC	2040
Db	1981	CACATCAATCTTACACCAAACTTGGAGCTTAATTTTCTACTATTCTCAGCAATAAC	2040
Qy	2041	ATTCTAAATATC 2052	
Db	2041	ATTCTAAAGATC 2052	
RESULT 3			
E40091			
LOCUS			
DEFINITION Plant promoter and terminator.			
ACCESSION E40091			
VERSION E40091.1 GI:18627207			
KEYWORDS JP 2000166577-A/5.			
SOURCE unidentified			
ORGANISM unclassified.			
REFERENCE 1. (bases 1 to 2056)			
AUTHORS Nishikawa,S. and Oeda,K.			
TITLE Plant promoter and terminator			
JOURNAL Patent: JP 2000166577-A 5 20-JUN-2000;			
COMMENT SUMITOMO CHEM CO LTD			
OS Daucus carota L.			
PN JP 2000166577-A/5			
PD 20-JUN-2000			
PF 01-OCT-1999 JP 1999281475			
PR			
PI SATOMI NISHIKAWA, KENJI OEDA			
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC			
C12N15/00,			
PC C12N5/00, (C12N5/00,C12R1:91)			
CC			
Key Location/Qualifiers			
FH promoter (1)-(2056).			
FT Location/Qualifiers			
1. 2056			
/organism="unidentified"			
/mol_type="genomic DNA"			
/db_xref="taxon:32644"			
FEATURES			
source			
ORIGIN			
Query Match 99.3%; Score 2038; DB 6; Length 2056;			
Best Local Similarity 99.8%; Pred. No. 6.7e-308;			
Matches 2052; Conservative 0; Mismatches 0; Indels 4; Gaps 1;			
Qy	1	CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTCTGA	60
Db	1	CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTCTGA	60
Qy	61	CTTCTCTCTTTTGACCTGTTGTATAAAGAAAGTAGAAATATTTTAAAAAGCTCGAA	120
Db	61	CTTCTCTCTTTTGACCTGTTGTATAAAGAAAGTAGAAATATTTTAAAAAGCTCGAA	120
Qy	121	TACTAACTTCTCTCAACAATCCGCTTCTTTTCCAAACACATTTATTAATCTTTTACT	180
Db	121	TACTAACTTCTCTCAACAATCCGCTTCTTTTCCAAACACATTTATTAATCTTTTACT	180

Qy	181	TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA	240
Db	181	TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA	240
Qy	241	AACGGCCTCAATAAAGATCATTTCAATGATATCTTTCAATTTTAGGATAACAATACGT	300
Db	241	AACGGCCTCAATAAAGATCATTTCAATGATATCTTTCAATTTTAGGATAACAATACGT	300
Qy	301	GAACAGGGTATTTTAACTGTCACCAAAATCTTAATAATTTTACCTGGCGGTGAACA	360
Db	301	GAACAGGGTATTTTAACTGTCACCAAAATCTTAATAATTTTACCTGGCGGTGAACA	360
Qy	361	CGGTCTTCCAGATAATATATTTTAAATTTGTAGCCTCCCTTTTAAACCAAAATTCGATGC	420
Db	361	CGGTCTTCCAGATAATATATTTTAAATTTGTAGCCTCCCTTTTAAACCAAAATTCGATGC	420
Qy	421	AGGACGACTTAGGTGAATACACATTTGTAAGTCTGAGTCTTTTAAACAAAGAAACAGTGGTTC	480
Db	421	AGGACGACTTAGGTGAATACACATTTGTAAGTCTTTTAAACAAAGAAACAGTGGTTC	480
Qy	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTACTATACTTT	540
Db	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTACTATACTTT	540
Qy	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATAATGCCCATCCAAAGGATAAGT	600
Db	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATAATGCCCATCCAAAGGATAAGT	600
Qy	601	AAAAATCCCGTTAAACCGAGTTGTTAAATAATATATGTTTACACTTACAGAGGATATTCGT	660
Db	601	AAAAATCCCGTTAAACCGAGTTGTTAAATAATATATGTTTACACTTACAGAGGATATTCGT	660
Qy	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAATGAGCGCTGTAACAGCCTAGACT	720
Db	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAATGAGCGCTGTAACAGCCTAGACT	720
Qy	721	TGGTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTAA	780
Db	721	TGGTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTAA	780
Qy	781	ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAGGATA	840
Db	781	ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAGGATA	840
Qy	841	AAAAAGAGCTTGTGCTGCTGTTAGTCTGTTGAGCTCATTTCTTTTAAAGTAATG	900
Db	841	AAAAAGAGCTTGTGCTGCTGTTAGTCTGTTGAGCTCATTTCTTTTAAAGTAATG	900
Qy	901	TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA	960
Db	901	TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA	960
Qy	961	TTAAACGAAAAATCAATTTTATAAATGCTCTCGGCTGCTATTATAATAGGATCACTTAC	1020
Db	961	TTAAACGAAAAATCAATTTTATAAATGCTCTCGGCTGCTATTATAATAGGATCACTTAC	1020
Qy	1021	TGATCATCTTAAACCTTTGTTTAAACCAAAATTTCAATGAGATAAAATATCTTTACAATGAA	1080
Db	1021	TGATCATCTTAAACCTTTGTTTAAACCAAAATTTCAATGAGATAAAATATCTTTACAATGAA	1080
Qy	1081	AAGAAGCAAAATGCTCTCTTTGAAAAAAACAAATAGGTACTCCCTCCCTCTGGAATGT	1140
Db	1081	AAGAAGCAAAATGCTCTCTTTGAAAAAAACAAATAGGTACTCCCTCCCTCTGGAATGT	1140
Qy	1141	ATACATATCGATTGGACACGGAGCTTAAGAAAAATGTTATAAAGTAATAGTAGTAATAAG	1200
Db	1141	ATACATATCGATTGGACACGGAGCTTAAGAAAAATGTTATAAAGTAATAGTAGTAATAAG	1200
Qy	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCACCAATATATAATTTGATAGATTAG	1260
Db	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCACCAATATATAATTTGATAGATTAG	1260

Qy	1321	AAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTG	1320
Db	1321	AAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAAAAATTTACTATTTTG	1320
Qy	1321	AGAAAGTTTTGAAATGTATAGAAATTGAGTGGGACATCCATAAAAGGAAAGTGATAGAAAT	1380
Db	1321	AGAAAGTTTTGAAATGTATAGAAATTGAGTGGGACATCCATAAAAGGAAAGTGATAGAAAT	1380
Qy	1381	TAAATGGGACAGAGGGAGTAAATACCTTTATGATATATAAAATTTTGTGTTATTTTGATTTCA	1440
Db	1381	TAAATGGGACAGAGGGAGTAAATACCTTTATGATATATAAAATTTTGTGTTATTTTGATTTCA	1440
Qy	1441	TAGATTATAAATCTATGTTTATATAGTAAATAATAATTTTAAAAATATACTATATTAAT	1500
Db	1441	TAGATTATAAATCTATGTTTATATAGTAAATAATAATTTTAAAAATATACTATATTAAT	1500
Qy	1501	CTGATTAGTCGATTCACGCCCTTTTATAAATTTTACAATACCTGAGTAAATATGAATAACAG	1560
Db	1501	CTGATTAGTCGATTCACGCCCTTTTATAAATTTTACAATACCTGAGTAAATATGAATAACAG	1560
Qy	1561	TTATCTGAAAGCAAAATAATATCTTTGTATAAAACAGCGTTTCGGTCAAAATGGGAAGTTCATG	1620
Db	1561	TTATCTGAAAGCAAAATAATATCTTTGTATAAAACAGCGTTTCGGTCAAAATGGGAAGTTCATG	1620
Qy	1621	TGTTATCCAAATAGTTTTTAATAATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA	1680
Db	1621	TGTTATCCAAATAGTTTTTAATAATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA	1680
Qy	1681	AATTTAAAAATAAATTTATGAGCATGGGAAGTTTACGGGCATCATGTGACGACACTAGACT	1740
Db	1681	AATTTAAAAATAAATTTATGAGCATGGGAAGTTTACGGGCATCATGTGACGACACTAGACT	1740
Qy	1741	GTTTGAACAATATGATGTCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT	1800
Db	1741	GTTTGAACAATATGATGTCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT	1800
Qy	1801	GCATT----CTGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTC	1856
Db	1801	GCATTCTAGTAGTAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTC	1860
Qy	1857	AACGGATTGGAATCCTTTTCTAAACTTTTAAAAATAAAAAATGCAATATGTAATATT	1916
Db	1861	AACGGATTGGAATCCTTTTCTAAACTTTTAAAAATAAAAAATGCAATATGTAATATT	1920
Qy	1917	TATCAACACCTCAACATGATGTTAGGTACTATAAATAGGTGCTCTTGGTGCTCTACTA	1976
Db	1921	TATCAACACCTCAACATGATGTTAGGTACTATAAATAGGTGCTCTTGGTGCTCTACTA	1980
Qy	1977	TCATCACAATCAATCTTACACCAAAACCTTTGAGCTTAATTTTCTACTTATTTCTCAGCAA	2036
Db	1981	TCATCACAATCAATCTTACACCAAAACCTTTGAGCTTAATTTTCTACTTATTTCTCAGCAA	2040
Qy	2037	TAACTTTCTAAATATC 2052	
Db	2041	TAACTTTCTAAATATC 2056	

RESULT	4
LOCUS	E40090
DEFINITION	E40090 Plant promoter and terminator.
ACCESSION	E40090
VERSION	E40090.1 GI:18627206
KEYWORDS	JP 2000166577-A/4.
SOURCE	unidentified
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 2048)
AUTHORS	Nishikawa, S. and Oeda, K.
TITLE	Plant promoter and terminator
JOURNAL	Patent: JP 2000166577-A 4 20-JUN-2000;
COMMENT	SWIMITOMO CHEM CO LTD OS Daucus carota L.

QY	781	ATTAGAGCTAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA	840
Db	777	ATTAGAGCTAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA	836
QY	841	AAACAGAGCTTGCTGCTGTGTGTTAGTTGTTGAGCTCAITTTCTTTTAAAGTAATG	900
Db	837	AAACAGAGCTTGCTGCTGTGTGTTAGTTGTTGAGCTCAITTTCTTTTAAAGTAATG	896
QY	901	TAACTGATCTAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATA	960
Db	897	TAACTGATCTAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATA	956
QY	961	TTAAACGAAATCATTTTATAACATGTCTCTCGGCTGTCTATTATAATAGGGATCACTTAC	1020
Db	957	TTAAACGAAATCATTTTATAACATGTCTCTCGGCTGTCTATTATAATAGGGATCACTTAC	1016
QY	1021	TGATCATCCATTAACCTTTGTTAAACAAATTCATATGAGATAAAATATCTTCAATGAA	1080
Db	1017	TGATCATCCATTAACCTTTGTTAAACAAATTCATATGAGATAAAATATCTTCAATGAA	1076
QY	1081	AAGAAGCAATGCTCTTTTGAAAAACAATAGGTACTCCCTCCGTCCTCTGAAATGT	1140
Db	1077	AAGAAGCAATGCTCTTTTGAAAAACAATAGGTACTCCCTCCGTCCTCTGAAATGT	1136
QY	1141	ATACATATGATTGGACACGAGACTAAGAAAAATGTATAAAGTAAATGTAGAGTAAAAAG	1200
Db	1137	ATACATATGATTGGACACGAGACTAAGAAAAATGTATAAAGTAAATGTAGAGTAAAAAG	1196
QY	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATATTTGATAGATTAG	1260
Db	1197	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATATTTGATAGATTAG	1256
QY	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAAATTTACTATTTTG	1320
Db	1257	AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAAATTTACTATTTTG	1316
QY	1321	AGAAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAGTGTATAGAAT	1380
Db	1317	AGAAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAGTGTATAGAAT	1376
QY	1381	TAAATGGGACAGGGAGTAAATACCTTTATGATATATAAATTTTGTATTTTGATTCTCA	1440
Db	1377	TAAATGGGACAGGGAGTAAATACCTTTATGATATATAAATTTTGTATTTTGATTCTCA	1436
QY	1441	TAGATTATAAATCTATGTTATAATGATATAATTTTAAAAATAATACATATATTAAT	1500
Db	1437	TAGATTATAAATCTATGTTATAATGATATAATTTTAAAAATAATACATATATTAAT	1496
QY	1501	CTGATTAGTCGATTACCGCCCTTTTATAATTTTACAATACCTAGTAAATATGAATAAATCAG	1560
Db	1497	CTGATTAGTCGATTACCGCCCTTTTATAATTTTACAATACCTAGTAAATATGAATAAATCAG	1556
QY	1561	TTATCTGAAAGCAAAATAATATCTTTGTAAGACAGCGTTCGGTCAAAATGGGAAGTTCAATG	1620
Db	1557	TTATCTGAAAGCAAAATAATATCTTTGTAAGACAGCGTTCGGTCAAAATGGGAAGTTCAATG	1616
QY	1621	TGTAATCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCGAGA	1680
Db	1617	TGTAATCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCGAGA	1676
QY	1681	AAATTAATAATTAATTTAGCATGGGAGCTTACCGGCATCATTTGAGCAGCACTAGACT	1740
Db	1677	AAATTAATAATTAATTTAGCATGGGAGCTTACCGGCATCATTTGAGCAGCACTAGACT	1736
QY	1741	GTTTGAACAATGTATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT	1800
Db	1737	GTTTGAACAATGTATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT	1796
QY	1801	GCATCTAGAAATACATCTTTTCAAAATTTCAACAAAACACAGCTTTTAACTTTTCTTCAACG	1860
Db	1797	GCATCTAGAAATACATCTTTTCAAAATTTCAACAAAACACAGCTTTTAACTTTTCTTCAACG	1856
QY	1861	GATTGGAATCCTTTTCTAAACCTTTTAAATAAAAAAATGCATTATTGTAATATTTATC	1920
Db	1857	GATTGGAATCCTTTTCTAAACCTTTTAAATAAAAAAATGCATTATTGTAATATTTATC	1916
QY	1921	AACACCTCAACATTGATGTTAGGCTACTATAATAATAGGTGCTCTTGGTGTCTACTATCAT	1980
Db	1917	AACACCTCAACATTGATGTTAGGCTACTATAATAATAGGTGCTCTTGGTGTCTACTATCAT	1976
QY	1981	CACATCAATCTTTACACACCAAAACCTTGAGCTTAAATTTTCTACTATTCTCAGCAATAAC	2040
Db	1977	CACATCAATCTTTACACACCAAAACCTTGAGCTTAAATTTTCTACTATTCTCAGCAATAAC	2036
QY	2041	ATTCTAAATATC 2052	
Db	2037	ATTCTAAATATC 2048	
RESULT 5			
E40089			
LOCUS			
DEFINITION			
E40089			
ACCESSION			
E40089.1			
VERSION			
JP 2000166577-A/3.			
KEYWORDS			
unidentified			
SOURCE			
unclassified			
ORGANISM			
1. (bases 1 to 2048)			
REFERENCE			
Nishikawa, S. and Oeda, K.			
AUTHORS			
Plant promoter and terminator			
TITLE			
Patent: JP 2000166577-A 3 20-JUN-2000;			
JOURNAL			
SUMITOMO CHEM CO LTD			
COMMENT			
OS Daucus carota L.			
PN JP 2000166577-A/3			
PD 20-JUN-2000			
PF 01-OCT-1999			
PR JP 1999281475			
PI SATOMI NISHIKAWA, KENJI OEDA			
PC C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC			
C12N15/00,			
PC C12N5/00, (C12N5/00, C12R1:91)			
CC			
FH Key			
FT promoter			
Location/Qualifiers			
1. .2048			
/organism="unidentified"			
/mol_type="genomic DNA"			
/db_xref="taxon:32644"			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2047; Conservative			
0; Mismatches			
1; Indels			
4; Gaps			
1;			
QY	1	CATGTGTCCCTACAGCACATAGGGGCTGTGTGTTGAGAGAGCAGAGCTGCTCTGA	60
Db	1	CATGTGTCCCTACAGCACATAGGGGCTGTGTGTTGAGAGAGCAGAGCTGCTCTGA	60
QY	61	CTTCTCTTCTTTTGACCTGTTTGTATAAAGAGTAGAATAATTTTAAAAAGCTCGAA	120
Db	61	CTTCTCTTCTTTTGACCTGTTTGTATAAAGAGTAGAATAATTTTAAAAAGCTCGAA	120
QY	121	TACTAACTCTCTCTCAACACTTCCGCTCTTTTCCAAACACCTTTTAACTTTTACT	180
Db	121	TACTAACTCTCTCTCAACACTTCCGCTCTTTTCCAAACACCTTTTAACTTTTACT	180
QY	181	TCCTATTCTTACTCCACTTCTTTTGGCTATAAGCAAGAAATCACCTCTTTTAAGCTAACCA	240
Db	181	TCCTATTCTTACTCCACTTCTTTTGGCTATAAGCAAGAAATCACCTCTTTTAAGCTAACCA	240
QY	241	AACGGCTCAATAAAAAAGATCAATCAATAATGTATCTTTCAATTTTAGGATAAACATACGT	300
Db			

1554 AAATCAGTATCTGAAAGCAATAATATCTTTGTAAACAGCG- - - - -TTCCGTCAAATG 1609
|||||
1543 AAATATTATCTGAATGATACATCTTTGTAAACAAACCTGGCCAAATAGACCATAA 1602
|||||
1610 GGAAGTTCATGTCATCAATAGTTTAAATATATAAAGTAAATTTTAAATTAATTTGTTATT 1669
|||||
1603 CCAAGTTCACGTGATCTTAAATGTTAAATACATAACATGAGTATTTCTTT - TTCAAGGT 1660
|||||
1670 TTTGTTTCAGAAATTTAAATAAATATTTAGAGCATGGAAAGTTTACCGGCATCATTTGAGC 1729
|||||
1661 ATAAGTTTAAATCTTCAATCAATTTAAATTTGAGACATTATTTGAGCAACTTTATGCC 1720
|||||
1730 AGCACTAGACTGTTTGAACAAATGATGTCGGGTGATCATCTATGACCTTTTCAACTCAAAC 1789
|||||
1721 CAGTTGATTTGTTAAACAAAGTGTGTCGGGTGATATTTATGACCTTTTCAACTCAAGC 1780
|||||
1790 TAGTGAAAT - AATGCAATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
|||||
1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACT 1840
|||||
1849 TTTCTTTCAACGGAATGGAATCTTTTCTAAATCTTTTAAATTTAAATAAATAATGATTTATT 1908
|||||
1841 TTTCTTTTAAACAGATTAGAAATCGTTTCTTAACTTTTAAATTT - AAAAAATACATTACT 1898
|||||
1909 GTAATATTATCAACACTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
|||||
1899 ATAATATTATCAACACTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958
|||||
1969 CTCTACTATCATCATCAATCTTTACACCAACCTTTGAGCTTAATTTTCTACTTTATT 2028
|||||
1959 CTCTACTATCATCATCAATCTTTCCAGCAACAACTTTGAGCTTAATTTTCTACTTAATT 2018
|||||
2029 CTGAGCAATCAATCTTCAATATC 2052
2019 TTTAGCAAAAACATCTTAAAGGTC 2042

RESULT 7
E15125
LOCUS E15125 2042 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15125
VERSION E15125.1 GI:5709808
KEYWORDS JP 1998052273-A/2.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 2 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/2
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FT source 1..2042 /organism='Daucus carota L.'
FT promoter 1<..<2042.
FT Location/Qualifiers
1..2042
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN

Query Match 11.8%; Score 241.2; DB 6; Length 2042;
Best Local Similarity 72.8%; Pred. No. 3.7e-28;
Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;
1554 AAATCAGTATCTGAAAGCAATAATATCTTTGTAAACAGCG- - - - -TTCCGTCAAATG 1609
|||||
1543 AAATATTATCTGAATGATACATCTTTGTAAACAAACCTGGCCAAATAGACCATAA 1602
|||||
1610 GGAAGTTCATGTCATCAATAGTTTAAATATATAAAGTAAATTTTAAATTAATTTGTTATT 1669
|||||
1603 CCAAGTTCACGTGATCTTAAATGTTAAATACATAACATGAGTATTTCTTT - TTCAAGGT 1660
|||||
1670 TTTGTTTCAGAAATTTAAATAAATATTTAGAGCATGGAAAGTTTACCGGCATCATTTGAGC 1729
|||||
1661 ATAAGTTTAAATCTTCAATCAATTTAAATTTGAGACATTATTTGAGCAACTTTATGCC 1720
|||||
1730 AGCACTAGACTGTTTGAACAAATGATGTCGGGTGATCATCTATGACCTTTTCAACTCAAAC 1789
|||||
1721 CAGTTGATTTGTTAAACAAAGTGTGTCGGGTGATATTTATGACCTTTTCAACTCAAGC 1780
|||||
1790 TAGTGAAAT - AATGCAATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
|||||
1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACT 1840
|||||
1849 TTTCTTTCAACGGAATGGAATCTTTTCTAAATCTTTTAAATTTAAATAAATAATGATTTATT 1908
|||||
1841 TTTCTTTTAAACAGATTAGAAATCGTTTCTTAACTTTTAAATTT - AAAAAATACATTACT 1898
|||||
1909 GTAATATTATCAACACTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
|||||
1899 ATAATATTATCAACACTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958
|||||
1969 CTCTACTATCATCATCAATCTTTACACCAACCTTTGAGCTTAATTTTCTACTTTATT 2028
|||||
1959 CTCTACTATCATCATCAATCTTTCCAGCAACAACTTTGAGCTTAATTTTCTACTTAATT 2018
|||||
2029 CTGAGCAATCAATCTTCAATATC 2052
2019 TTTAGCAAAAACATCTTAAAGGTC 2042

RESULT 8

AR076816
LOCUS AR076816 247 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 1 from patent US 5959176.
ACCESSION AR076816
VERSION AR076816.1 GI:10003562
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 1 28-SEP-1999;
FEATURES
source 1..247
/organism='unknown'
/mol_type='unassigned DNA'

ORIGIN

Query Match 9.8%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.9e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
1804 TTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACCGAT 1863
|||||
1 TTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTTCTTTTAAACAGAT 60
|||||
1864 TGGAATCCCTTTTCTTAACTTTTAAATTTAAATAAATAATGCAATTATTTGTAATTTATCAAC 1923
|||||


```
Db 61 TAGAATCGTTCTCAAACTTTTAAATTT--AAAAATACATTACTATAATATTATCAAC 118
Qy 1924 ACCTCAACATGTAGTGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCAC 1983
Db 119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCAC 178
Qy 1984 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATACATT 2043
Db 179 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTAATTTTCTAGCAAAACATT 238
Qy 2044 CTAATATC 2052
Db 239 CTAAGGTC 247

RESULT 9
LOCUS E15124
DEFINITION Promoter.
ACCESSION E15124
VERSION E15124.1 GI:5709807
KEYWORDS JP 1998052273-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 1 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/1
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
scrandness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..247
FT /organism="Daucus carota L."
FT /clones="pCR16GI-Xb"
FT promoter 1<..247.
FT Location/Qualifiers
FEATURES
source 1..247
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 9.6%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.9e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy 1804 TTCTAGAATCATCTTTTCAAAATTTCAACAACACAGCTTTAACTTTTCTTCAACGGAT 1863
Db 1 TTCTAGAATATATCTTTTGAAATTTCAACAACACAGCACTAATTTCTTTTAACAGAT 60
Qy 1864 TGAATCCTTTCTTAAACTTTTAAAAATAAAAAATGCAATTTGTAATATTATCAAC 1923
Db 61 TAGAATCGTTCTCAAACTTTTAAATTT--AAAAATACATTACTATAATATTATCAAC 118
Qy 1924 ACCTCAACATGTAGTGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCAC 1983
Db 119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCAC 178
Qy 1984 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATACATT 2043
Db 179 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTAATTTTCTAGCAAAACATT 238
```

```
Qy 2044 CTAATATC 2052
Db 239 CTAAGGTC 247

RESULT 10
LOCUS ARI146852
DEFINITION Sequence 2 from patent US 6218598.
ACCESSION ARI146852
VERSION ARI146852.1 GI:15110041
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,F., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: US 6218598-A 2 17-APR-2001;
FEATURES Location/Qualifiers
source 1..246
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy 1805 TCTAGAATCATCTTTTCAAAATTTCAACAACACAGCTTTAACTTTTCTTCAACGGATT 1864
Db 1 TCTAGAATATATCTTTTGAAATTTCAACAACACAGCACTAATTTCTTTTAACAGATT 60
Qy 1865 GGAATCCTTTTCTAAACTTTTAAAAATAAAAAATGCAATTTGTAATATTATCAACA 1924
Db 61 AGAATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACTATAATATTATCAACA 118
Qy 1925 CCTCAACATGTAGTGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCAC 1984
Db 119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCAC 178
Qy 1985 TCAATCTTACACCACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATACATT 2044
Db 179 TCAATCTTCCAGCAAAACCTTGAGCTTAATTTTCTACTAATTTTCTAGCAAAACATT 238
Qy 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 11
LOCUS E55065
DEFINITION Plant promoter.
ACCESSION E55065
VERSION E55065.1 GI:18625251
KEYWORDS JP 2000083679-A/2.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,I., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: JP 2000083679-A 2 28-MAR-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000083679-A/2
PD 28-MAR-2000
PF 12-JUL-1999 JP 1999197240
PR IKUHARU ISHIGE, SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//((C12N15/09,C12R1:91), PC
(C12N5/10,C12R1:91), C12N15/00,C12N5/00,C12R1:91), PC
```

(C12N5/00, C12R1.91)	
CC	Key
FT	source
FT	Location/Qualifiers
FT	1..246
FT	/organism="Daucus carota L."
FEATURES	
source	1..246
	/organism="unidentified"
	/mol_type="genomic DNA"
	/db_xref="taxon:32644"
ORIGIN	
Query Match 9.5%; Score 195.6; DB 6; Length 246;	
Best Local Similarity 89.5%; Pred. No. 7e-21;	
Matches 22; Conservative 0; Mismatches 24; Indels 2; Gaps 1;	
Qy	1805 TCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTTCAACGATT 1864
Db	1 TCTAGATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTTCTTTTAAACGATT 60
Qy	1865 GGAATCCTTTTCTAACTTTTAAATAAAAAAATGCATTATTGTAATATTTTATCAACA 1924
Db	61 AGAATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACTATTAATATTTATCAACA 118
Qy	1925 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCACCA 1984
Db	119 CCTCAACATTCACTGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCACCA 178
Qy	1985 TCAATCTTACACCAACACCTTGAGCTTAATTTTCTTACTTATTTCTCAGCAATAACATTC 2044
Db	179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTTCAGCAAAACATTC 238
Qy	2045 TAAATATC 2052
Db	239 TAAAGGTC 246
RESULT 12	
DCA18706/c	
LOCUS	DCA18706 8005 bp DNA linear PLN 10-FEB-1999
DEFINITION	Daucus carota Inv*Dc5 gene.
ACCESSION	Y18706
VERSION	Y18706.1 GI:4454114
KEYWORDS	beta-fructofuranosidase; Inv*Dc5 gene; invertase; isoform.
SOURCE	Daucus carota (carrot)
ORGANISM	Daucus carota
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae; Daucus.
REFERENCE	1
AUTHORS	Sturm, A.
TITLE	Molecular characterisation and functional analysis of sucrose-cleaving enzymes in carrot (Daucus carota L.)
JOURNAL	J. Exp. Bot. 47, 1187-1192 (1996)
REFERENCE	2 (bases 1 to 8005)
AUTHORS	Sturm, A.
TITLE	Direct Submission
JOURNAL	Submitted (27-JAN-1999) A. Sturm, Friedrich Miescher-Institut, Maulbeestr. 66, CH-4058 Basel, SWITZERLAND
COMMENT	Related sequence X67163.
FEATURES	Location/Qualifiers
source	1..8005
	/organism="Daucus carota"
	/mol_type="genomic DNA"
	/cultivar="Nantaise"
	/db_xref="taxon:4039"
	1..8005
	/genes="Inv*Dc5"
gene	1..2666
5'UTR	/genes="Inv*Dc5"
CDS	join(2767..3160,3724..3732,4656..5515,5605..5766,5869..6104,6190..6277,6384..6587)
/gene="Inv*Dc5"	
/EC_number="3.2.1.26"	
/note="soluble acid invertase"	
/codon_start=1	
/product="beta-fructofuranosidase, isoform II"	
/protein_id="CAA77266.1"	
/db_xref="GI:4454115"	
/db_xref="GOA:Q42722"	
/db_xref="UniProt/TREMBL:Q42722"	
/translation="MEHPITISHYTEPLDGEHSPSLTTNTAEQSSRRSLTFLVLLFS	
SILAACLVMGTWVLPNSGNEAVEKSTVVFETVEVAPRGVAEGVSMKSRPALNAE	
PPANFPWNSNLISWQSSPHFQPNQNMNDPNGPLFYKGYHLFYQNPDPGAIGNKI	
VMGHAVSSDLIIHWKHLPLVAMVTDHWYDVANGVTGSATILPDGQIVMLYTGSTNESQV	
LAKTAYPADSPDLLIEWKYPGNPVLVPPGIDFKDRDPTTAMRTPEGKRWLIIGSK	
QNLGGLSLAVYDVTDFKNFTLLDQVLHAVHGTGMWECVDFVPVSKFGENGLDTSFGDVG	
VXHWKASLDDDRNDYYAIGTYDVPVSGKWPNPNELDVGILGYDYGIIYASKTFYDS	
NKKRWLWIKETDSEISDRKGMASVQGIPTKILFDPKTGNLNLQWPEEVKNKURL	
NKTVFVNINTGAVLPLEIGSGQLDITAEFEVDKESLERVQETNEVYDCKNGGSS	
GRGALGPFGLLIADKDLSEQTFVYFIAGKSGNLRITFFCADHSRSSKAVDYDKGIY	
GSVVPVLRGEKLTMRILVDHSIVESFSQGRITCITSRVYPTKAIYNNAKVFLFNNAE	
ARIIASLNTWQMTAQRTQTHFADLVI"	
exon	2767..3160
	/gene="Inv*Dc5"
	/number=1
intron	3161..3723
	/gene="Inv*Dc5"
	/number=1
exon	3724..3732
	/genes="Inv*Dc5"
	/number=2
intron	3733..4655
	/genes="Inv*Dc5"
	/number=2
exon	4656..5515
	/genes="Inv*Dc5"
	/number=3
intron	5516..5604
	/gene="Inv*Dc5"
	/number=3
exon	5605..5766
	/genes="Inv*Dc5"
	/number=4
intron	5767..5868
	/gene="Inv*Dc5"
	/number=4
exon	5869..6104
	/genes="Inv*Dc5"
	/number=5
intron	6105..6189
	/genes="Inv*Dc5"
	/number=5
exon	6190..6277
	/gene="Inv*Dc5"
	/number=6
intron	6278..6383
	/genes="Inv*Dc5"
	/number=6
exon	6384..6587
	/genes="Inv*Dc5"
	/number=7
3'UTR	6588..8005
	/gene="Inv*Dc5"
ORIGIN	
Query Match 8.5%; Score 174; DB 8; Length 8005;	
Best Local Similarity 81.7%; Pred. No. 8.2e-18;	
Matches 201; Conservative 0; Mismatches 45; Indels 0; Gaps 0;	
Qy	17 CACATAGGCGCTGTTTGGTTGAGAGAGCAGAGCTGCTCTCTCTCTCTCTTTGA 76
Db	1285 CACTATGGGGGTGTTTGTTCGCCGAAGTAGGAGCTGCTTCGCTCTCTCTCTGA 1226
Qy	77 CCTGTTTGTATAAAGAAGTAGAATATTTTAAAAAGCTGCGAATACTACTACTCTCTC 136

```
Db 1225 CCGGTTGTGTAATAAGCAGACATTTTAAAAAGTTGAGATGCTAGCTTCTCTC 1166
Qy 137 ACAATTCGCTCTCTTTTCCAAACACATTTTAACTTTTACCTTCTCTCTCTCTCTCA 196
Db 1165 ACAGCTTCTGCTCTTTTCCAAACACATTTTAACTTTTACCTTCTCTCTCTCTCTCA 1106
Qy 197 CTTCTTTGCTATAGCAAGAAATCCTCTTTTAACTTACCTTCTCTCTCTCTCTCTCTCA 256
Db 1105 CTTTCTTAATTAAGTAAGAAATCCTCTTTTAACTTACCTTCTCTCTCTCTCTCTCTCA 1046
Qy 257 GATCAT 262
Db 1045 GACCT 1040

RESULT 13
BD188672
LOCUS BD188672 2831 bp DNA linear PAT 17-JUL-2003
DEFINITION Promotor and use thereof.
ACCESSION BD188672
VERSION BD188672.1 GI:32998411
KEYWORDS JP 2003000252-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2831)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Promotor and use thereof.
JOURNAL Patent: JP 2003000252-A 1 07-JAN-2003;
SUMITOMO CHEMICAL CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 2003000252-A/1
PD 07-JAN-2003
PF 31-MAY-2001 JP 2001164069
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/19,C12N1/21,C12N5/10,C12N15/
CC 00,C12N5/00
FH Promotor and use thereof
FT promoter Location/Qualifiers
1. .2831
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 7.9%; Score 162.4; DB 6; Length 2831;
Best Local Similarity 82.5%; Pred. No. 6.4e-16;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
Qy 22 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTCTCTCTCTCTCTCTCT 81
Db 1716 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTCTCTCTCTCTCTCTCT 1774
Qy 82 TTGTATAAGAGATAGAAATATTTTAAAGCTGCGAATCTCTCTCTCTCTCTCTCTCTCAAC 141
Db 1775 TTGTATAAGAGATAGAAATATTTTAAAGCTGCGAATCTCTCTCTCTCTCTCTCTCTCAAC 1834
Qy 142 TTGCGCTTCTTTTCCAAACACATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTCT 201
Db 1835 TTGCGCTTCTTTTCCAAACACATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTCT 1894
Qy 202 TTGCTATAAGCAAGAAATCTCTTTTAAAGCTAACCCAAACGGCTCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAAATCTCTTTTAAAGCTAACCCAAACGGCTCTCAATAAAGATCA 1954

RESULT 14
BD188674
LOCUS BD188674 2865 bp DNA linear PAT 17-JUL-2003
DEFINITION Promotor and use thereof.
```

```
ACCESSION BD188674
VERSION BD188674.1 GI:32998413
KEYWORDS JP 2003000252-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2865)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Promotor and use thereof
JOURNAL Patent: JP 2003000252-A 3 07-JAN-2003;
SUMITOMO CHEMICAL CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 2003000252-A/3
PD 07-JAN-2003
PF 31-MAY-2001 JP 2001164069
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/
CC 00,C12N5/00
FH Promotor and use thereof
FT promoter Location/Qualifiers
1. .2865
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 7.9%; Score 162.4; DB 6; Length 2865;
Best Local Similarity 82.5%; Pred. No. 6.4e-16;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
Qy 22 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTCTCTCTCTCTCTCTCT 81
Db 1716 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTCTCTCTCTCTCTCTCT 1774
Qy 82 TTGTATAAGAGATAGAAATATTTTAAAGCTGCGAATCTCTCTCTCTCTCTCTCTCTCAAC 141
Db 1775 TTGTATAAGAGATAGAAATATTTTAAAGCTGCGAATCTCTCTCTCTCTCTCTCTCTCAAC 1834
Qy 142 TTGCGCTTCTTTTCCAAACACATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTCT 201
Db 1835 TTGCGCTTCTTTTCCAAACACATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTCT 1894
Qy 202 TTGCTATAAGCAAGAAATCTCTTTTAAAGCTAACCCAAACGGCTCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAGTCAATCTTTTAAATTAACCCAAACGGCCCTTAAGTAATTGA 1954

RESULT 15
DARGCHS2
LOCUS DARGCHS2 4886 bp DNA linear PLN 14-APR-2000
DEFINITION Carrot gCHS2 gene for chalcone synthase.
ACCESSION D16255
VERSION D16255.1 GI:441168
KEYWORDS chalcone synthase.
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae;
Daucinae; Daucus.
REFERENCE 1 (bases 1 to 4886)
AUTHORS Ozeki,Y., Davies,E. and Takeda,J.
TITLE Structure and expression of chalcone synthase gene in carrot
JOURNAL Plant Cell Physiol. 34, 1029-1037 (1993)
REFERENCE 2 (bases 1 to 4886)
AUTHORS Ozeki,Y.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1993) Yoshihiro Ozeki, College of Arts and
Sciences, The University of Tokyo, Department of Biology, Komaba,
Meguro-ku, Tokyo 153, Japan (Tel:03-3467-1171 (ex.253)),
```

COMMENT Fax:03-3485-2904)
Submitted (14-MAY-1993) to DDBJ by:
Yoshihiro Ozeki
Department of Biol. College of
Arts & Science, Univ. of Tokyo
3-8-1 Komaba, Meguro-ku
Tokyo 153
Japan

Phone: 03-3467-1171 x253

Fax: 03-3485-2904.

Location/Qualifiers

1. .4886

/organism="Daucus carota"

/mol_type="genomic DNA"

/cultivar="Kurodagozun"

/db_xref="taxon:4039"

/dev_stage="cultured cells"

1957. .1962

/note="TACPyAT motif"

misc_feature

misc_feature

misc_feature

misc_feature

TATA_signal

prim_transcript

exon

gene

CDS

intron

exon

polyA_signal

ORIGIN

Query Match

Best Local Similarity

Matches

198; Conservative

0; Mismatches

41; Indels

1; Gaps

1;

11 CTACAGCACATAGGGCTGTTGGTTCAGAGAGCAGAGCTGCTTCTGACTTCTTCTTC 70

1355 CTAGTAGATCTAGGGCTGTTGGATGTTGGAGAGCAGAGCTGCTTCCGGCTTCTGCTTT 1414

71 TTTTGTACCTCTTTGTATAAGAGAGTAAATATTTTAAAGAGCTGCGAATACCTTAACTTC 130

1415 TCTTGACCTCTTTGGTAAAGAGTAGAGCCTTTTAAAGAGTTAGAGATAGTACTAGCTTC 1474

131 TCTCTCAAACTTCG-CTTCTTTTCAAAACATTTTAACTTTTAACTTTTCTCAATTC 189

1475 TCTCTCAGAGCTTCTGTTTTTTTCCAAACATTTTATTATTATTATTATTATTATTATT 1534

190 TACTCCACTCTTCTGTATTAAGCAAGAGTACCTTCTTTAAAGCTAACCCAAACGGCTC 249

1535 TACTCCACTCTTCTTAATTAAGCAAGAGTACCTTCTTTTAAACTAACCCAAACGGCTC 1594

Search completed: March 14, 2005, 18:28:48
Job time : 9015 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 14:16:29 ; Search time 1121 Seconds
(without alignments)
10836.137 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 cagtgtgcctacagcaca.....gcaataacattctaaatatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2052	100.0	2052	3 AAA37959	Aaa37959 Carrot pr
2	2048.8	99.8	2052	3 AAA37961	Aaa37961 Carrot pr
3	2034.8	99.2	2056	3 AAA37964	Aaa37964 Plasmid #
4	2030.8	99.0	2048	3 AAA37962	Aaa37962 Plasmid #
5	2030.8	99.0	2048	3 AAA37963	Aaa37963 Plasmid #
6	239.6	11.7	2042	2 AAV15144	AAV15144 New promo
7	196.6	9.6	247	2 AAV15143	AAV15143 New promo
8	195.6	9.5	246	3 AAZ49611	Aaz49611 Carrot CR
9	194.4	9.5	196	12 ADP07499	Adp07499 Carrot DN
10	162.4	7.9	2831	10 ADC56759	Adc56759 Carrot DN
11	162.4	7.9	2865	10 ADC56761	Adc56761 Carrot DN
c 12	119	5.8	140	3 AAZ49616	Aaz49616 Oligonucle
13	119	5.8	140	3 AAZ49615	Aaz49615 Oligonucle
14	99.4	4.8	8056	8 ABZ10246	Abz10246 Haematopo
15	91.8	4.5	8056	8 ABZ10100	Abz10100 Haematopo
c 16	91.6	4.5	8056	8 ABZ10246	Abz10246 Haematopo
c 17	88.6	4.3	851	3 AAA37960	Aaa37960 Carrot te
c 18	88.6	4.3	8056	8 ABZ10100	Abz10100 Haematopo
c 19	87.2	4.2	110000	13 ABD32968_6	Continuation (7 of
20	83	4.0	158001	12 ADL17884	Adl17884 Human pho

c 21	75	3.7	11745	6 ABK28332	Abk28332 DNA trans
c 22	71.2	3.5	8222	8 ACF62816	Acf62816 Colon can
c 23	71.2	3.5	11222	10 ADB54318	Adb54318 Pretreat
c 24	71	3.5	5930	6 ABL32517	AbL32517 Human imm
c 25	70.8	3.5	5286	13 ADS89278	Ads89278 Oligonucle
c 26	70.8	3.5	5286	13 ADS89552	Ads89552 Oligonucle
c 27	69.8	3.4	6352	6 ABK31340	Abk31340 Signal tr
c 28	69.8	3.4	6352	6 ABL70563	AbL70563 Chemicall
c 29	69.8	3.4	6352	6 AAS61235	Aas61235 Human gen
c 30	69.6	3.4	8222	8 ACF62794	Acf62794 Colon can
c 31	69.6	3.4	11222	10 ADB54190	Adb54190 Pretreat
c 32	69.4	3.4	16258	6 ABK40038	Abk40038 Human che
c 33	69.4	3.4	16258	6 ABL70376	AbL70376 Chemicall
c 34	69.2	3.4	6045	6 ABK31541	Abk31541 Signal tr
c 35	69.2	3.4	6045	6 ABL70624	AbL70624 Chemicall
c 36	68.2	3.3	15548	6 ABL34155	AbL34155 Human imm
c 37	68.2	3.3	110000	13 ABD32968_6	Continuation (7 of
c 38	68	3.3	11422	6 ABK39936	Abk39936 Human che
c 39	68	3.3	11422	6 ABL32218	AbL32218 Human imm
c 40	67.8	3.3	15732	4 AAS45388	Aas45388 Chemicall
c 41	67.8	3.3	15732	6 ABK28233	Abk28233 DNA trans
c 42	67.8	3.3	99764	13 ABD32992	Abd32992 Human can
c 43	67.6	3.3	5984	6 ABQ66994	Abq66994 Human ang
c 44	67.6	3.3	6050	6 ABL34011	AbL34011 Human imm
c 45	67.4	3.3	5286	13 ADS89551	Ads89551 Oligonucle

ALIGNMENTS

RESULT 1

AAA37959

ID AAA37959 standard; DNA; 2052 BP.

XX

AC AAA37959;

XX

DT 18-AUG-2000 (first entry)

XX

DE Carrot promoter sequence #1.

XX

KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.

XX

OS Daucus carota.

XX

WO2000020613-A1.

PN

PD 13-APR-2000.

XX

PF 28-SEP-1999; 99WO-JP005303.

XX

PR 02-OCT-1998; 98JP-00281124.

XX

PA (SUMO) SUMITOMO CHEM CO LTD.

XX

PI Nishikawa S, Oeda K;

XX

DR WPI; 2000-303791/26.

XX

PT New Plant promoters and terminators from Daucus carota L., useful in

PT

PR plant breeding, for e.g. controlling fertilities of plants.

PS Claim 1; Page 69-70; 81pp; English.

XX

CC This sequence represents a carrot promoter. The invention relates to

CC plant promoters and terminators from Daucus carota L. which are capable

CC of expressing a gene of interest in plants. The invention also includes a

CC chimeric gene characterized in that it comprises the promoter and a

CC desired gene linked to each other in the form capable of functioning. A

CC method of producing a transformant comprises introducing the promoter,

CC the chimeric gene or a vector comprising the promoter and a desired gene

CC or terminator sequence into a host cell. The plant promoters and

CC terminators are useful in plant breeding, for e.g. fertilities of plants

CC may be controlled by expressing, in the host cells, a sense or antisense

CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;
Query Match 100.0%; Score 2052; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGTGCGCCCTACGACACATAGGCGCTGTTGGTTGAGAGAGCGAGACGCTCTCTGA 60
DB 1 CATGTGCGCCCTACGACACATAGGCGCTGTTGGTTGAGAGAGCGAGACGCTCTCTGA 60
QY 61 CTTCTCTCTCTTTGACCTGTTGTATAAAGAAAGTAGAAATATTTTAAAGCTCGAA 120
DB 61 CTTCTCTCTCTTTGACCTGTTGTATAAAGAAAGTAGAAATATTTTAAAGCTCGAA 120
QY 121 TACTAATCTCTCTCAAACTTCCGCTTCTTTTCCAAACACCTTTATTAATCTTTTACT 180
DB 121 TACTAATCTCTCTCAAACTTCCGCTTCTTTTCCAAACACCTTTATTAATCTTTTACT 180
QY 181 TCTCATTTCTACTCCACTCTTTGCTATAAGCAAGAAATCACTTCTTTAAGCTAACCCA 240
DB 181 TCTCATTTCTACTCCACTCTTTGCTATAAGCAAGAAATCACTTCTTTAAGCTAACCCA 240
QY 241 AACGGCTCAATAAAGATCAATTAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
DB 241 AACGGCTCAATAAAGATCAATTAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
QY 301 GAACAGGGTTATTTTAAACGCTGTCAACAAATTTCTAATTAATTTTACCTGCGGTGAACA 360
DB 301 GAACAGGGTTATTTTAAACGCTGTCAACAAATTTCTAATTAATTTTACCTGCGGTGAACA 360
QY 361 CCGTCTTCCAAGATAATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGATGC 420
DB 361 CCGTCTTCCAAGATAATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGATGC 420
QY 421 AGACGACTTAGGTGAATACACATGTGACGTCTTTTAAACAAAGAACAGTGGTTC 480
DB 421 AGACGACTTAGGTGAATACACATGTGACGTCTTTTAAACAAAGAACAGTGGTTC 480
QY 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACCAACCTCTATCCAGCTACTACTTT 540
DB 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACCAACCTCTATCCAGCTACTACTTT 540
QY 541 TGGCCGAATCGTCTCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAAGT 600
DB 541 TGGCCGAATCGTCTCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAAGT 600
QY 601 AAAATCCCGTTTAAACAGTTTGTATATATATGTTTACATTTACAAGAGGATATTCGT 660
DB 601 AAAATCCCGTTTAAACAGTTTGTATATATATGTTTACATTTACAAGAGGATATTCGT 660
QY 661 AATACCTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
DB 661 AATACCTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
QY 721 TGGTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATAAA 780
DB 721 TGGTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATAAA 780
QY 781 ATTAGAGCTAATTAATTAAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
DB 781 ATTAGAGCTAATTAATTAAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
QY 841 AAAACAAGAGCTTCTGCTGTGTGTTTGTGTTGAGCTCATTTCTTTAAAAGTAATG 900
DB 841 AAAACAAGAGCTTCTGCTGTGTGTTTGTGTTGAGCTCATTTCTTTAAAAGTAATG 900
QY 901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA 960
DB 901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA 960
QY 961 TTAACGAAAATCAITTTTATAACATGTCTCTCGCTGCTCATTTATTAATAGGATCACTTAC 1020

DB 961 TTAACGAAAATCAITTTTATAACATGTCTCTCGCTGCTCATTTATTAATAGGATCACTTAC 1020
QY 1021 TGATCATCCATTAAACCTTTGTTAAACAAATTCATGAGATAAAATATCTTACAATGAA 1080
DB 1021 TGATCATCCATTAAACCTTTGTTAAACAAATTCATGAGATAAAATATCTTACAATGAA 1080
QY 1081 AAGAAGGACAAATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCTGAAATGT 1140
DB 1081 AAGAAGGACAAATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCTGAAATGT 1140
QY 1141 ATACATATCGAATTGGACACGAGACTAAGAAAAATGTATATAAGTAATPAGAGTAAAAAG 1200
DB 1141 ATACATATCGAATTGGACACGAGACTAAGAAAAATGTATATAAGTAATPAGAGTAAAAAG 1200
QY 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATATAATTTGATAGATTTAG 1260
DB 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATATAATTTGATAGATTTAG 1260
QY 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGATTTTTATATATAAAAAATTTACTATTTTG 1320
DB 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGATTTTTATATATAAAAAATTTACTATTTTG 1320
QY 1321 AGAAAGTTTTGAAATGTATAGAAATTGAGTGGGACATCCATAAAGGAAAAAGTGTATAGAAT 1380
DB 1321 AGAAAGTTTTGAAATGTATAGAAATTGAGTGGGACATCCATAAAGGAAAAAGTGTATAGAAT 1380
QY 1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTGATTCTCA 1440
DB 1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTGATTCTCA 1440
QY 1441 TAAGATTATAAATCTATGTTTATAATGATAATATAATTTTAAAAAATACTATATTAAT 1500
DB 1441 TAAGATTATAAATCTATGTTTATAATGATAATATAATTTTAAAAAATACTATATTAAT 1500
QY 1501 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATACTGAGTAAATGAATAAATCAG 1560
DB 1501 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATACTGAGTAAATGAATAAATCAG 1560
QY 1561 TTATCTGAAAAAGCAATAATATCTTTGTGTAACAGCGTTCCGTCATAATGGGAAGTTCATG 1620
DB 1561 TTATCTGAAAAAGCAATAATATCTTTGTGTAACAGCGTTCCGTCATAATGGGAAGTTCATG 1620
QY 1621 TGTATTCAATAGCTTTTAAATATAAAGTAATTTTAAATTAATTTTGTATTTTGTTCGATCAGA 1680
DB 1621 TGTATTCAATAGCTTTTAAATATAAAGTAATTTTAAATTAATTTTGTATTTTGTTCGATCAGA 1680
QY 1681 AATTTAAAAATAAATTAATTGAGCATGGGAAAGTTTACCGGCAATCAATTGAGCAGCACTAGACT 1740
DB 1681 AATTTAAAAATAAATTAATTGAGCATGGGAAAGTTTACCGGCAATCAATTGAGCAGCACTAGACT 1740
QY 1741 GTTTGAAACAATGTATGTCGGTGTACATCTATGACCTTTTCAACTCAACTCAACTAGTGAATAAT 1800
DB 1741 GTTTGAAACAATGTATGTCGGTGTACATCTATGACCTTTTCAACTCAACTCAACTAGTGAATAAT 1800
QY 1801 GCATTTCTAGAATACATCTTTTCAAAATTTTCAAAACAACACAGCTTTTAACTTTTCAACG 1860
DB 1801 GCATTTCTAGAATACATCTTTTCAAAATTTTCAAAACAACACAGCTTTTAACTTTTCAACG 1860
QY 1861 GATTGGAATCCTTTTCTTAAACTTTTAAAAATAAAAAAATGCAATTAATTTGTAATTTATC 1920
DB 1861 GATTGGAATCCTTTTCTTAAACTTTTAAAAATAAAAAAATGCAATTAATTTGTAATTTATC 1920
QY 1921 AACACCTCAACATGTATGTTAGCGTACTATATAAGTGTCTTTGGTGTCTACTATCAT 1980
DB 1921 AACACCTCAACATGTATGTTAGCGTACTATATAAGTGTCTTTGGTGTCTACTATCAT 1980
QY 1981 CACATCAATCTTACACCAACCTTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2040
DB 1981 CACATCAATCTTACACCAACCTTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2040
QY 2041 ATTCTAAATATC 2052
|||||

Db	2041 ATTCTAAATATC 2052	Qy	301	GAAACAGGGTTATTTTTTAAACGCTGTCACAAATTTCTAATAATTTTACCTGGCGGCTGAACA	360
Db		Qy	301	GAAACAGGGTTATTTTTTAAACGCTGTCACAAATTTCTAATAATTTTACCTGGCGGCTGAACA <th>360</th>	360
Db		Qy	361	CGTCTTCCAGAGTAATAATTTTTTAAATTTTGTAGGCTCCCTTTTAAACCAAAATTCGCATGC <th>420</th>	420
Db		Qy	361	CGTCTTCCAGAGTAATAATTTTTTAAATTTTGTAGGCTCCCTTTTAAACCAAAATTCGCATGC <th>420</th>	420
Qy		Qy	421	AGGACGACTTAGGTGAATAACATTTGTACTGTGAGTCTTTTAAACCAAAAGCAAGTGGTTC <th>480</th>	480
Db		Qy	421	AGGACGACTTAGGTGAATAACATTTGTACTGTGAGTCTTTTAAACCAAAAGCAAGTGGTTC <th>480</th>	480
Qy		Qy	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTAATTTTACCTTT <th>540</th>	540
Db		Qy	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTAATTTTACCTTT <th>540</th>	540
Qy		Qy	541	TGSCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATAATGCCCCATCCAAAGGATAAGT <th>600</th>	600
Db		Qy	541	TGSCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATAATGCCCCATCCAAAGGATAAGT <th>600</th>	600
Qy		Qy	601	AAAAATCCCGTTTAAACGAGTTTGTAAATATATATATTTTACACTTCAAGAGGATAATTCGT <th>660</th>	660
Db		Qy	601	AAAAATCCCGTTTAAACGAGTTTGTAAATATATATATTTTACACTTCAAGAGGATAATTCGT <th>660</th>	660
Qy		Qy	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAATTTGACGCTGGTAAACGCTTAGACT <th>720</th>	720
Db		Qy	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAATTTGACGCTGGTAAACGCTTAGACT <th>720</th>	720
Qy		Qy	721	TGCTCACTGATTAATAGATAATTTGTAGTATAATAATAGTACGATCTACAATGACATTTAA <th>780</th>	780
Db		Qy	721	TGCTCACTGATTAATAGATAATTTGTAGTATAATAATAGTACGATCTACAATGACATTTAA <th>780</th>	780
Qy		Qy	781	ATTAGAGCTATTAAATTAAGTTACTAAATAAAGAGAGGTTAGTAAACAGAAAGCAGGTA <th>840</th>	840
Db		Qy	781	ATTAGAGCTATTAAATTAAGTTACTAAATAAAGAGAGGTTAGTAAACAGAAAGCAGGTA <th>840</th>	840
Qy		Qy	841	AAAAAAGAGCTTGCTGCTGCTGTTTGTAGTTGTTGTAGCTCATTTCTTTTAAAGTAATG <th>900</th>	900
Db		Qy	841	AAAAAAGAGCTTGCTGCTGCTGTTTGTAGTTGTTGTAGCTCATTTCTTTTAAAGTAATG <th>900</th>	900
Qy		Qy	901	TAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA <th>960</th>	960
Db		Qy	901	TAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA <th>960</th>	960
Qy		Qy	961	TTAAACGAAAATCATTTTATAACATGCTCTCCGCTGTCATTATAATAGGATCACTTAC <th>1020</th>	1020
Db		Qy	961	TTAAACGAAAATCATTTTATAACATGCTCTCCGCTGTCATTATAATAGGATCACTTAC <th>1020</th>	1020
Qy		Qy	1021	TGATCATCCATTTAAACACCTTGTAAAAACAAATTTCAATGAGATAAAATATCTTTACAAATG <th>1080</th>	1080
Db		Qy	1021	TGATCATCCATTTAAACACCTTGTAAAAACAAATTTCAATGAGATAAAATATCTTTACAAATG <th>1080</th>	1080
Qy		Qy	1081	AAGAAGCAATGCTCTCTTTTGAAGAAACAAATAGTACTCCCTCCGCTCCCTCTGAAATGT <th>1140</th>	1140
Db		Qy	1081	AAGAAGCAATGCTCTCTTTTGAAGAAACAAATAGTACTCCCTCCGCTCCCTCTGAAATGT <th>1140</th>	1140
Qy		Qy	1141	ATACATATGATTTGGACACGAGAGCTAAGAAAAATGTTAAGTAATCTAGAGTAAAAAG <th>1200</th>	1200
Db		Qy	1141	ATACATATGATTTGGACACGAGAGCTAAGAAAAATGTTAAGTAATCTAGAGTAAAAAG <th>1200</th>	1200
Qy		Qy	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCCAATATATAATTTTGTATAGATTTAG <th>1260</th>	1260
Db		Qy	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCCAATATATAATTTTGTATAGATTTAG <th>1260</th>	1260
Qy		Qy	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTTTATATATATAAAAAATTTTACTATTTTG <th>1320</th>	1320
Db		Qy	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTTTATATATAAAAAATTTTACTATTTTG <th>1320</th>	1320
Qy		Qy	1321	AGAAAGTTTGAATGATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAT <th>1380</th>	1380
Db		Qy	1321	AGAAAGTTTGAATGATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAT <th>1380</th>	1380

Matches 2046; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

Qy	1	CATGTGCGCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAAGAGCTGCTTCTGA	60
Db	1	CATGTGCGCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAAGAGCTGCTTCTGA	60
Qy	61	CTTCTCTCTCTTTTGACCTGTTTGTATAAAGAAGTAGAATAATTTTAAAGAGCTCGGAA	120
Db	61	CTTCTCTCTCTTTTGACCTGTTTGTATAAAGAAGTAGAATAATTTTAAAGAGCTCGGAA	120
Qy	121	TACTAACTTCTCTCTCACAACTCCGCTTCTTTTCCAAACACATTTATTAACTTTTTTACT	180
Db	121	TACTAACTTCTCTCTCACAACTCCGCTTCTTTTCCAAACACATTTATTAACTTTTTTACT	180
Qy	181	TCTCATTTTCTACCCACTTCTTGCTATAAGCAAGAAATCACATCTTTTAAAGCTAACCCA	240
Db	181	TCTCATTTTCTACCCACTTCTTGCTATAAGCAAGAAATCACATCTTTTAAAGCTAACCCA	240
Qy	241	AACGGCTCAATAAAGATCATTTCAATAATGTATCTTCAATTTTAAAGGATAACAAATACGT	300
Db	241	AACGGCTCAATAAAGATCATTTCAATAATGTATCTTCAATTTTAAAGGATAACAAATACGT	300
Qy	301	GAACAGGGTTATTTTTTAAAGTGTCAACAAATTTCTAATAATTTTAACTGCGCGGTGAACA	360
Db	301	GAACAGGGTTATTTTTTAAAGTGTCAACAAATTTCTAATAATTTTAACTGCGCGGTGAACA	360
Qy	361	CGCTCTTCCAAGATATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGATGC	420
Db	361	CGCTCTTCCAAGATATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGATGC	420
Qy	421	AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAGAACAGTGGTTC	480
Db	421	AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAGAACAGTGGTTC	480
Qy	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCACGTAATATCTTT	540
Db	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCACGTAATATCTTT	540
Qy	541	TGGCGAATGCTCTCAAAATGTTTTTATATGTATTAATAATGCCCATCCAGGATAAGT	600
Db	541	TGGCGAATGCTCTCAAAATGTTTTTATATGTATTAATAATGCCCATCCAGGATAAGT	600
Qy	601	AAATATCCCGTTTAAACAGTTTGTATATATATGTTTACATTTACAAGAGGATATTCGT	660
Db	601	AAATATCCCGTTTAAACAGTTTGTATATATATGTTTACATTTACAAGAGGATATTCGT	660
Qy	661	AATACTTTTACGACAGAGACTTAGGTCAAAATGGACGCTGGTAAACAGCCTAGACT	720
Db	661	AATACTTTTACGACAGAGACTTAGGTCAAAATGGACGCTGGTAAACAGCCTAGACT	720
Qy	721	TGGTCACTGATAATAGATTAATTTGTAGTATATATAGTAGGATCTACAAATGATTA	780
Db	721	TGGTCACTGATAATAGATTAATTTGTAGTATATATAGTAGGATCTACAAATGATTA	780
Qy	781	ATTAGAGCTAATTAATTAAGTTACTTAATAAAGAGAGGTTAGTAAACAGAAAGAGGTA	840
Db	781	ATTAGAGCTAATTAATTAAGTTACTTAATAAAGAGAGGTTAGTAAACAGAAAGAGGTA	840
Qy	841	AAACCAAGAGCTTCTGCTGTGTGTGTAGTTGTGTGAGCTCAATTTTAAAGTAATG	900
Db	841	AAACCAAGAGCTTCTGCTGTGTGTGTAGTTGTGTGAGCTCAATTTTAAAGTAATG	900
Qy	901	TAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA	960
Db	901	TAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA	960
Qy	961	TTAAACGAAATCAATTTTATAACATGTCTCTCGCTGTCTATATAATAGGATCACTTAC	1020
Db	961	TTAAACGAAATCAATTTTATAACATGTCTCTCGCTGTCTATATAATAGGATCACTTAC	1020
Qy	1021	TGATCATCCATTAACACCTTTGTTAAACAAATTCATGAGATAAAATATCTTCAATGAA	1080
Db	1017	TGATCATCCATTAACACCTTTGTTAAACAAATTCATGAGATAAAATATCTTCAATGAA	1076

RESULT 5
AAA37963
ID

AAA37963 standard; DNA; 2048 BP.

Db 1437 TAAGATTATAAATCTATGTTTATAATGATAATAAATTTTAAAAAATAACTATATTAAT 1496
Qy 1501 CTGATTAGTCGATTACCGCCCTTTTATAATTTTCAATACACTGAGTATATGAATAAATCAG 1560
Db 1497 CTGATTAGTCGATTACCGCCCTTTTATAATTTTCAATACACTGAGTATATGAATAAATCAG 1556
Qy 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1620
Db 1557 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1616
Qy 1621 TGTATTCAATAGTTTAAATAATAAAGTAAATTTTAAATTAATTTTGTATTTTGTTCAGA 1680
Db 1617 TGTATTCAATAGTTTAAATAATAAAGTAAATTTTAAATTAATTTTGTATTTTGTTCAGA 1676
Qy 1681 AATTTAAATAAATATTGAGCATGGGAAGTTTCAAGGCAATCATTTGAGCAGCACTAGACT 1740
Db 1677 AATTTAAATAAATATTGAGCATGGGAAGTTTCAAGGCAATCATTTGAGCAGCACTAGACT 1736
Qy 1741 GTTTGAACAATGATGTCGGGTGATCATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db 1737 GTTTGAACAATGATGTCGGGTGATCATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1796
Qy 1801 GCATTCTAGATAATCATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG 1860
Db 1797 GCATTCTAGATAATCATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG 1856
Qy 1861 GATTGGAATCCTTTTCAAACTTTTAAATAAATAAATAAATGCAATTTTGTATAATTTATC 1920
Db 1857 GATTGGAATCCTTTTCAAACTTTTAAATAAATAAATAAATGCAATTTTGTATAATTTATC 1916
Qy 1921 AACACCTCAACATTTGATGTTAGCGTACTATAAATAGTGTCTCTTGGTGTCTACTATCAT 1980
Db 1917 AACACCTCAACATTTGATGTTAGCGTACTATAAATAGTGTCTCTTGGTGTCTACTATCAT 1976
Qy 1981 CATCATCAATCTTACACCAAACTTTGAGCTTAAATTTTCTACTATTTCTCAGCAATAAC 2040
Db 1977 CATCATCAATCTTACACCAAACTTTGAGCTTAAATTTTCTACTATTTCTCAGCAATCAG 2036
Qy 2041 ATTCTAAATATC 2052
Db 2037 ATTCTAAAGATC 2048

RESULT 6
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX
AC AAV15144;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX WPI; 1998-122310/12.
XX

XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 2; Page 15-16; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;
Query Match 11.7%; Score 239.6; DB 2; Length 2042;
Best Local Similarity 72.6%; Pred. No. 8.7e-35;
Matches 366; Conservative 0; Mismatches 129; Indels 9; Gaps 4;
Qy 1554 AATCAGTTTATCTGAAAGCAAAATAATATCTTTGTAAACACAGCG---TTCCGTCAAATG 1609
Db 1543 AATATTATTATCTGAATGATAAATCATCTTTGTAAACAAAACTGCGCAATAGGACCATAA 1602
Qy 1610 GGAAGTTTCATGTTATTTCAATAGTTTAAATAAAGTAAATTTTAAATTAATTTGTTATT 1669
Db 1603 CCNAGTTTCAGTGTATTTTAAATGTTAATACTAATGAGTATTTTCTT--TTCAAGGT 1660
Qy 1670 TTTGTTTCAGAAATTTAAATAAATTTATGAGCATGGGAAGTTTCAGGGCATCATTTGAGC 1729
Db 1661 ATAAGTTAATCTTCAATCAATTAATTTTAAATTTGGACATTTATGAGCAACTTTATGCC 1720
Qy 1730 AGCCTAGACTCTTTCAACAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAAC 1789
Db 1721 CAGTTTGTATTTTAAACACAGCTTTGTCGGGTGATATTTATGACCTTTCAACTCAAG 1780
Qy 1790 TAGTGAAT--AATGCAATTTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
Db 1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACT 1840
Qy 1849 TTTCTTTTCAACGGATGGAAATCCTTTTCTAAATCTTTTAAATTAATAAATAAATGCAATATT 1908
Db 1841 TTTCTTTTAAACAGATTAGAATCGTTTTCGTAAACCTTTTAAATTT--AAAAATACATTACT 1898
Qy 1909 GTAATATTTTATCAACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
Db 1899 ATAATATTTTATCAACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958
Qy 1969 CTCTACTATCATCATCAATCTTACACCAAACTTCAAGCAAACTTGAAGCTTAATTTTCTACTATT 2028
Db 1959 CTCTACTATCATCATCAATCTTCCAGCAAACTTGAAGCTTAATTTTCTACTATT 2018
Qy 2029 CTCAGCAATAACATTCTAAATATC 2052
Db 2019 TTTAGCAAAACATTTCTAAAGGTC 2042
RESULT 7
AAV15143
ID AAV15143 standard; DNA; 247 BP.
XX
AC AAV15143;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX


```
PF 18-OCT-2002; 2002JP-00304115.
XX
PR 18-OCT-2002; 2002JP-00304115.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2004-360986/34.
XX
XX Expression inducing promoter, useful for expressing foreign gene,
PT comprising first DNA for detecting transcription start point and having
PT minimum promoter function, linked to second DNA having expression
PT inducing promoter function.
XX
PS Disclosure; SEQ ID NO 4; 61pp; Japanese.
XX
CC The invention relates to an expression inducing promoter comprising a
CC first DNA linked to a second DNA at its 5' terminus, where the first DNA
CC has a region which determines a transcription start point of RNA
CC polymerase II and has minimum promoter function, and the second DNA has
CC expression inducing promoter function in a plant cell. The expression
CC inducing promoter is useful for expressing a foreign gene. This sequence
CC represents carrot DNA used in the scope of the invention.
XX
SQ Sequence 196 BP; 66 A; 40 C; 19 G; 71 T; 0 U; 0 Other;
Query Match 9.5%; Score 194.4; DB 12; Length 196;
Best Local Similarity 99.5%; Pred. No. 1.4e-26;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1853 TTTCACGGATTGGAATCCTTTCTAAACCTTTTAAATAAAAAAATGCAATTATTGTAA 1912
DB 1 TTTCACGGATTGGAATCCTTTCTAAACCTTTTAAATAAAAAAATGCAATTATTGTAA 60
QY 1913 TATTTATCAACACCTCAACATTGATGTAGCGTACTATAAATAGGTGCTCTTGGTGCTCT 1972
DB 61 TATTTATCAACACCTCAACATTGATGTAGCGTACTATAAATAGGTGCTCTTGGTGCTCT 120
QY 1973 ACTATCATCATCATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATTCTCA 2032
DB 121 ACTATCATCATCATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATTCTCA 180
QY 2033 GCAATACATTCTAAA 2048
DB 181 GCAATACATTCTAAA 196
RESULT 10
ADC56759
ID ADC56759 standard; DNA; 2831 BP.
XX
AC ADC56759;
XX
DT 18-DEC-2003 (first entry)
XX
DE Carrot DNA that encodes an expression inducer type promoter.
XX
KW carrot; expression inducer type promoter; plant; vector; transformant;
XX ds.
XX
OS Daucus carota subsp. sativus.
XX
PN JP2003000252-A.
XX
XX 07-JAN-2003.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX WPI; 2003-472920/45.
XX
XX
XX This invention relates to carrot DNA used as an expression inducer type
XX promoter. Specifically, it refers to promoter sequences derived from
XX plant DNA, preferably carrot, and functional mutants thereof that can
XX be
```

```
PT An expression inducer type promoter derived from DNA of Daucus carota L.
PT var. sativa DC. of 2831 bases and its analogues having the same
XX functions.
PS Claim 1; SEQ ID NO 1; 16pp; Japanese.
XX
CC This invention relates to carrot DNA used as an expression inducer type
CC promoter. Specifically, it refers to promoter sequences derived from
CC plant DNA, preferably carrot, and functional mutants thereof that can be
CC used as expression inducers. Furthermore, the present invention describes
CC DNA, vectors, transformants and the process by which to prepare
CC transformants. This polynucleotide sequence is the DNA encoding the
CC carrot promoter of the invention.
XX
SQ Sequence 2831 BP; 950 A; 456 C; 395 G; 1030 T; 0 U; 0 Other;
Query Match 7.9%; Score 162.4; DB 10; Length 2831;
Best Local Similarity 82.5%; Pred. No. 1.5e-20;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 22 AGGGCTGTTTGGTTGAGAGAGAGAGAGAGCTTCTGACTTCTTCTTCTTTGACTGT 81
DB 1716 AGGGCTGTTTGGTTTATGGAATCAGAGCTGCTTCTGACTTCTTCTTCTTTT-ACCCGT 1774
QY 82 TTGTATAAAGAGTAGAATATTTTAAAGCTGCGAATACTACTTCTCTCACAAC 141
DB 1775 TTGTGTAAAAAGAGAGAGAGAGAGCTTTTAAAGAGCTGAGATGCTAGTTTCTCTCAGC 1834
QY 142 TTCGCTCTTTTCCAAACACTTTATTAACCTTTTAACTTTTCTTCATTTCTTCTTCTTCTTCT 201
DB 1835 TTCTGCTCTTTTCCAAACACTTTATCACTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1894
QY 202 TTGCTATAAGCAAGAAATCACTTCTTTAAGCTAACCCAGGCTCAATAAAGATCA 261
DB 1895 TTACTATAAGCAAGAGTCAATTTCTTTTAAATTAACCCAAAGCCCTTAAAGTAATTGA 1954
RESULT 11
ADC56761
ID ADC56761 standard; DNA; 2865 BP.
XX
AC ADC56761;
XX
DT 18-DEC-2003 (first entry)
XX
DE Carrot DNA encoding an expression inducer type promoter (SeqID 3).
XX
KW carrot; expression inducer type promoter; plant; vector; transformant;
XX ds.
XX
OS Daucus carota subsp. sativus.
XX
PN JP2003000252-A.
XX
XX 07-JAN-2003.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX WPI; 2003-472920/45.
XX
XX
XX An expression inducer type promoter derived from DNA of Daucus carota L.
XX var. sativa DC. of 2831 bases and its analogues having the same
XX functions.
PS Disclosure; SEQ ID NO 3; 16pp; Japanese.
XX
XX This invention relates to carrot DNA used as an expression inducer type
XX promoter. Specifically, it refers to promoter sequences derived from
XX plant DNA, preferably carrot, and functional mutants thereof that can
XX be
```


CC used as expression inducers. Furthermore, the present invention describes
 CC DNA, vectors, transformants and the process by which to prepare
 CC transformants. This polynucleotide sequence is DNA encoding a carrot
 CC promoter (SeqID 3) of the invention.

XX SQ Sequence 2865 BP; 963 A; 461 C; 396 G; 1045 T; 0 U; 0 Other;

Query Match 7.9%; Score 162.4; DB 10; Length 2865;
 Best Local Similarity 82.5%; Pred. No. 1.5e-20;
 Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Qy 22 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTTCTTTGACCTGT 81

Db 1716 AGGCGCTGTTGGTTTATGGAATCAGAAGCTGCTTCTGACTTCTTCTTCTTTT-ACCGGT 1774

Qy 82 TTGTATTAAGAGTAGAATAATTTTTTAAAGAGCTGCGAATACTTCTCTCTCAACAAC 141

Db 1775 TTGTGTAATAAGAGCAGAGCACTTTTAAGAAGCTGAGAATGCTAGTTCTCTCTCACAGC 1834

Qy 142 TTCGCTTCTTTTCCAAACACTTTTATTAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 201

Db 1835 TTCTGCTTCTTTTCCAAACACTTTTATCAACTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1894

Qy 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCTCAATAAAGATCA 261

Db 1895 TTACTATAAGCAAGAGTCAATTCTTTTAAATTAACCCAAACGGCCCTTAAGTAATTGA 1954

RESULT 12

AZ49616/c
 ID AZ49616 standard; DNA; 140 BP.

XX AC AZ49616;

XX DT 07-APR-2000 (first entry)

XX DE Oligonucleotide-4 for synthesis of CR16.1 fragment for plant promoter.

XX KW Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
 KW soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
 KW male sterility-related gene; ss.

XX OS Daucus carota.

XX OS Synthetic.

XX PN EP976832-A2.

XX PD 02-FEB-2000.

XX PF 13-JUL-1999; 99EP-00113732.

XX PR 15-JUL-1998; 98JP-00200372.

XX PA (SUMO) SUMITOMO CHEM CO LTD.

XX PI Ishige F, Nishikawa S, Oeda K;

XX WPI; 2000-128374/12.

XX PT Novel promoter used to produce transgenic plants with higher expression
 PT of a desired gene.

XX PS Disclosure; Page 14; 24pp; English.

XX CC The present sequence is an oligonucleotide (- chain) used to prepare a
 CC CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
 CC fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
 CC construction of a plant promoter. The promoter is used for controlling
 CC the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
 CC desaturase and S-locus type specific RNase gene (male sterility-related
 CC gene) in a host cell especially a microorganism or a plant cell. The
 CC transformed plant cells can be used to produce transgenic plants. The
 CC promoter is compact and therefore suitable for higher expression of a

CC desired gene in a particular tissue compared to other host tissues
 XX SQ Sequence 140 BP; 45 A; 17 C; 34 G; 44 T; 0 U; 0 Other;

Query Match 5.8%; Score 119; DB 3; Length 140;
 Best Local Similarity 92.6%; Pred. No. 1.1e-12;
 Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1918 ATCAACACCTCAACATTGATGTTAGCGTACTATATAAATAGGTGCTCTTGGTCTCTACTAT 1977

Db 140 ATCAACACCTCAACATTGATGTTAGCGTACTATATAAATAGGTGCTCTTGGTCTCTACTAT 81

Qy 1978 CATCACATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAAT 2037

Db 80 CATCACATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTTATTTTCTAGCAAA 21

Qy 2038 AACATTTCTAAATATC 2052

Db 20 AACATTTCTAAAGTTC 6

RESULT 13

AZ49615
 ID AZ49615 standard; DNA; 140 BP.

XX AC AZ49615;

XX DT 07-APR-2000 (first entry)

XX DE Oligonucleotide-3 for synthesis of CR16.1 fragment for plant promoter.

XX KW Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
 KW soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
 KW male sterility-related gene; ss.

XX OS Daucus carota.

XX OS Synthetic.

XX PN EP976832-A2.

XX PD 02-FEB-2000.

XX PF 13-JUL-1999; 99EP-00113732.

XX PR 15-JUL-1998; 98JP-00200372.

XX PA (SUMO) SUMITOMO CHEM CO LTD.

XX PI Ishige F, Nishikawa S, Oeda K;

XX WPI; 2000-128374/12.

XX PT Novel promoter used to produce transgenic plants with higher expression
 PT of a desired gene.

XX PS Disclosure; Page 13; 24pp; English.

XX CC The present sequence is an oligonucleotide (+ chain) used to prepare a
 CC CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
 CC fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
 CC construction of a plant promoter. The promoter is used for controlling
 CC the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
 CC desaturase and S-locus type specific RNase gene (male sterility-related
 CC gene) in a host cell especially a microorganism or a plant cell. The
 CC transformed plant cells can be used to produce transgenic plants. The
 CC promoter is compact and therefore suitable for higher expression of a
 CC desired gene in a particular tissue compared to other host tissues

SQ Sequence 140 BP; 44 A; 34 C; 17 G; 45 T; 0 U; 0 Other;

Query Match 5.8%; Score 119; DB 3; Length 140;
 Best Local Similarity 92.6%; Pred. No. 1.1e-12;
 Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1502 TGATTAGTCGATACCGCTTTTATATATTTTACAACTACTGAGTAATATGAATAAATCAGT 1561
 Db 2926 AATTTTATGAAATGTAATAATTTAAATGAATATATTTGTTTAAATAAATAAATAAATAATTA 2985
 QY 1562 TATCTGAAAGCAAAATATATCTTTTGTAAACACAGCGTTCGGTCAAAATGGGAAGTTCATGT 1621
 Db 2986 TTGAAAAATTTGATATATTTTAAATGATGATGTTTAAATTTTATAATATAAATAAATAA 3045
 QY 1622 GTATTCAAATGTTTATATATAAAGTAATAATTTTAAATTAATATGTTTATTTTTCGAA 1681
 Db 3046 TAAATAAATGTTTATATATGTAATGTAATGAAATTTGAAATTTATTTTATTTTATGTTAA 3105
 QY 1682 ATTTAAATAAATATTTGAGCATGGGAGTTTCAGGCATCATTTGACGACACTAGACTG 1741
 Db 3106 AATAAAAAATTAATTTTAAATGAAATTTAAATGATGATGAAATTTGTTTATATAT 3165
 QY 1742 ---TTTGAACAATGTATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATA 1798
 Db 3166 AATTTTAAATAAATAATTTTAAATAAATTTTGTGTTGTAATGTAATAAATAATTT 3225
 QY 1799 ATGCATTTCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAA 1858
 Db 3226 AATTTATATATTTTAAATAAATTTGATTTATTTTAAATTAATGATGTAATAAATAA 3285
 QY 1859 CGAATGGAATCCTTTTCTAAACTTTTAAATAAATAAATAAATGCAATTTGTAATATTA 1918
 Db 3286 AATTTTGTGTTATTTGTTGTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3345
 QY 1919 TCAACACCTCAACATTCATGTTA 1941
 Db 3346 AAAATGAATAAATTTGAAATGA 3368

RESULT 15

ID ABZ10100 standard; DNA; 8056 BP.
 XX ABZ10100;

DT 16-JAN-2003 (first entry)

XX Haematopoietic cell proliferation disorder related DNA sequence #240.

XX Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.

XX Homo sapiens.

XX WO200277272-A2.

XX 03-OCT-2002.

XX 26-MAR-2002; 2002WO-EP003401.

XX 26-MAR-2001; 2001US-0278333P.

XX (EPIG-) EPIGENOMICS AG.

XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwöpe I, Ziebarth H;

XX WPI; 2003-018942/01.

XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.

PS Claim 28; SEQ ID NO 240; 117pp; English.

XX The present invention describes a method for detecting and

CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB211118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 XX

SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;

Query Match 4.5%; Score 91.8; DB 8; Length 8056;
 Best Local Similarity 44.2%; Pred. No. 1.8e-07;
 Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5;

QY 547 AATGCTTCTCAAAATGTTTTTATATGTAATAAATGCCCATCCACAGGATAAGTAAATTT 606
 Db 1358 ATTTTATTAATAAATTTATTTTATATAAATAATATCGATTTTATTTTATTTT 1417
 QY 607 CCCGTTTAAACAGTTTGTATATATATATGTTTACACTTACAGAGGATATTCGTAATCT 666
 Db 1418 TTTATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1477
 QY 667 TTTAGACGACAGAGCTTAGTCAAAATGGAGCGTGTAAACAGCCTAGACTGTGCTCA 726
 Db 1478 ATAAATTAATTAATAATTAATATATATAAATAAATAAATAAATAAATAAATAAATAA 1537
 QY 727 CTGATAAATAGATAATTTGTAGTATATAATAGTAGGATCTACATGACATTAATAATAGA 786
 Db 1538 TTAATAAATAATTTTAAAAAATAATATTTAAATTTAAATTTAAATTTAAATTTAAATTT 1597
 QY 787 GCTATTAAATTAAGTTACTATAATAAAGAGAGGTTAGTAAACAGAAAGCGGTAAAAACA 846
 Db 1598 TAATTTTAAATTTTAAAAAATCGAAATAACGAATCGTAAATAAATAAATAAATAAATAA 1657
 QY 847 AGAGCTTGTGCTGTGTTAGTTAGTTGTGAGCTCATTTCTTTAAAGTAATGTAATCT 906
 Db 1658 AATTAATAC-GATAAAAAATTTTATTTTATTAATAAATAAATAAATAAATAAATAAATAA 1716
 QY 907 GATCTAAAGCACATAGAAATTTAGTACAGCTTAAACCTTTTACAGAATTTATATTAAC 966
 Db 1717 TATATATTTAATAATTTTAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAA 1776
 QY 967 GAAAAATCAATTTTATAACATGCTCTCGGCTGTCTATTATAATAGGATCACTTACTGATCA 1026
 Db 1777 AATTAATAAATAATTAATGCTTATATATTTTAAATTAATAAATAAATAAATAAATAAATAA 1836
 QY 1027 TCCATTAAAAACCTTTTAAACCAAAATCAATGAGATATAATATCTTCAATGAAAGAGAG 1086
 Db 1837 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1896
 QY 1087 GACAATGCTCTTTTGAATAAACAATAAGTACTCCCTCCCTCTCGCTGTAATGATACAT 1146
 Db 1897 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1956
 QY 1147 ATGGATTGGACACGAGACTAAGAAAAATGTAATAAGTATAGTAGTAAAAAGAGAGAG 1206
 Db 1957 AATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2016
 QY 1207 A-----AAGAAAAGTGGGTAAAGTAGCGGGACCCCAATATATATTAATGATAGAT 1256

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 15:39:49 ; Search time 355 Seconds
(without alignments)
9458.140 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcctacagcaca.....gcataacattctaataatc 2052

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241.2	11.8	2042	US-08-911-434A-2	Sequence 2, Appli
2	196.6	9.6	247	US-08-911-434A-1	Sequence 1, Appli
3	195.6	9.5	246	US-09-352-608-2	Sequence 2, Appli
4	119	5.8	140	US-09-352-608-6	Sequence 6, Appli
5	119	5.8	140	US-09-352-608-7	Sequence 7, Appli
6	81.4	4.0	187169	US-09-949-016-12776	Sequence 12776, A
7	81.4	4.0	181569	US-09-949-016-15940	Sequence 15940, A
8	72.2	3.5	205044	US-09-949-016-15851	Sequence 15851, A
9	72.2	3.5	205044	US-09-949-016-15852	Sequence 15852, A
10	72.2	3.5	205044	US-09-949-016-15853	Sequence 15853, A
11	72.2	3.5	223471	US-09-949-016-12387	Sequence 12387, A
12	72.2	3.5	223471	US-09-949-016-12724	Sequence 12724, A
13	72.2	3.5	223471	US-09-949-016-12725	Sequence 12725, A
14	69.4	3.4	7218	US-08-232-463-14	Sequence 14, Appl
15	68.6	3.3	1141	US-09-806-708B-22	Sequence 22, Appl
16	68.6	3.3	18773	US-09-949-016-14164	Sequence 14164, A
17	68.2	3.3	1141	US-09-806-708B-22	Sequence 22, Appl
18	67.8	3.3	205044	US-09-949-016-15851	Sequence 15851, A
19	67.8	3.3	205044	US-09-949-016-15852	Sequence 15852, A
20	67.8	3.3	205044	US-09-949-016-15853	Sequence 15853, A
21	67.8	3.3	223471	US-09-949-016-12387	Sequence 12387, A
22	67.8	3.3	223471	US-09-949-016-12724	Sequence 12724, A
23	67.8	3.3	223471	US-09-949-016-12725	Sequence 12725, A
24	67.2	3.3	601	US-09-949-016-30530	Sequence 30530, A
25	67.2	3.3	601	US-09-949-016-30531	Sequence 30531, A
26	67.2	3.3	601	US-09-949-016-37149	Sequence 37149, A
27	67.2	3.3	601	US-09-949-016-37150	Sequence 37150, A

Query Match 11.8%; Score 241.2; DB 2; Length 2042;
Best Local Similarity 72.8%; Pred. No. 5.4e-42;

ALIGNMENTS

RESULT 1

US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satoshi
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,434A
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..2042
; US-08-911-434A-2

Sequence 37163, A
Sequence 37164, A
Sequence 145867,
Sequence 145868,
Sequence 146135,
Sequence 146136,
Sequence 146403,
Sequence 146404,
Sequence 30531, A
Sequence 37150, A
Sequence 37164, A
Sequence 145868,
Sequence 146136,
Sequence 146404,
Sequence 156535,
Sequence 16110, A
Sequence 11934, A
Sequence 13248, A

Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;	
QY 1554	AAATCAGTATCTGAAAGCAATAATATCTTTGTAACACAGCG----TTCCGTCAAATG 1609
Db	
QY 1543	AAATTTATCTGAATGATAACATCTTTGTAACAAACCTGGCCAAATAGCACCATAA 1602
Db	
QY 1610	GGAAGTTCATGTCATTCATAGTTTAAATATAAAGTAAATTTTAAATTAATTTGTTAT 1669
Db	
QY 1603	CCAAGTTCAGGTGATTTCTTAAATGTTAACTAAACATGAGTATTTCTT--TTCAAGGT 1660
Db	
QY 1670	TTTGTTCAGAAATTTAAATAAATTTAGCATGGGAAGTTTCAGGGCATCAITTGAGC 1729
Db	
QY 1661	ATAAGTTAATCTTCAATCAATTAACCTTTTAAATTTGGACATTAATGAGCAACTTTATGCC 1720
Db	
QY 1730	AGCACTAGACTGTTTGAACAAATGATGTCGGGTGATCATCTATGACCTTTTCAACTCAAAC 1789
Db	
QY 1721	CAGTGTATGTTTAAACAAAGTGTGTCGGGTGATTAATTAATGACCTTTTCAACTCAAGC 1780
Db	
QY 1790	TAGTGAAT-AATGCATTTCTAGAATACATCTTTTCAAAATTTTCAACAAACAGAGCTTTAACT 1848
Db	
QY 1849	TTTCTTTCAACGGAATGGAATCCTTTTCTAAACCTTTTAAATAAATAAAATGCATTAAT 1908
Db	
QY 1841	TTTCTTTTAAACAGATAGAAATCGTTTCTTAAACCTTTTAAATTT--AAAAAATACATTA 1898
Db	
QY 1909	GTAATATTTTCAACACCTCAACATGATGTAGCGTACTATAAATAGGTGCTCTGGTG 1968
Db	
QY 1899	ATATATTTTATCAACACCTCAACATCATGTAGCGTACTATAAATAGGTGCTCTGGTG 1958
Db	
QY 1969	CTCTACTATCATCATCAATCTTTTACACCAAAACCTTTGAGCTTAAATTTTCTACTTAAT 2028
Db	
QY 1959	CTCTACTATCATCATCAATCTTTTCCAGCAGCAAAACCTTTGAGCTTAAATTTTCTACTTAAT 2018
Db	
QY 2029	CTCAGCAATCAATCTTCAATATC 2052
Db	
QY 2019	TTTAGCAAAACAACTTCTAAAGTC 2042
Db	
RESULT 2 911-434A-1	
US-08-911-434A-1	
; Sequence 1, Application US/08911434A	
; Patent No. 5959176	
; GENERAL INFORMATION:	
; APPLICANT: TORIKAI, Satomi	
; APPLICANT: OEDA, Kenji	
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF	
; NUMBER OF SEQUENCES: 14	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP	
; STREET: P.O. BOX 747	
; CITY: FALLS CHURCH	
; STATE: VIRGINIA	
; COUNTRY: UNITED STATES OF AMERICA	
; ZIP: 22040	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PC-DOS/MS-DOS	
; SOFTWARE: PatentIn Release #1.0, Version #1.30	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/08/911,434A	
; FILING DATE: 12-AUG-1997	
; CLASSIFICATION: 800	
; ATTORNEY/AGENT INFORMATION:	
; NAME: Stewart, Raymond C.	
; REGISTRATION NUMBER: 21,066	
; REFERENCE/DOCKET NUMBER: 2185-0199P	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: (703)205-8000	
; TELEFAX: (703)205-8050	
; INFORMATION FOR SEQ ID NO: 1:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 247 base pairs	
; TYPE: nucleic acid	
; STRANDEDNESS: single	
; TOPOLOGY: linear	
; MOLECULE TYPE: DNA (genomic)	
; HYPOTHETICAL: NO	
; ORIGINAL SOURCE:	
; ORGANISM: Daucus carota L.	
; INDIVIDUAL ISOLATE: Kuroda Gosun	
; FEATURE:	
; NAME/KEY: promoter	
; LOCATION: 1..247	
; US-08-911-434A-1	
Query Match 9.6%; Score 196.6; DB 2; Length 247;	
Best Local Similarity 89.6%; Pred. No. 1e-32;	
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;	
QY 1804	TTCTAGAATACATCTTTTCAAAATTTCAACAAACAGCTTTTAACTTTCTTTCAACGGAT 1863
Db	
QY 1864	TGGAATCCTTTTCTAAACCTTTTAAATAAATAAATAATGCATTAATTTATCAAC 1923
Db	
QY 1924	ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCAC 1983
Db	
QY 119	ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCAC 178
Db	
QY 1984	ATCAATCTTTTACACCAAAACCTTGAGCTTAAATTTTCTACTTAATTTCTCAGCAATAACAT 2043
Db	
QY 179	ATCAATCTTTCCAGCAGCAAAACCTTGAGCTTAAATTTTCTACTTAATTTTAGCAAAAACAT 238
Db	
QY 2044	CTAAATATC 2052
Db	
QY 239	CTAAAGGTC 247
Db	
RESULT 3	
US-09-352-608-2	
; Sequence 2, Application US/09352608	
; Patent No. 6218598	
; GENERAL INFORMATION:	
; APPLICANT: ISHIGE, Fumiharu	
; APPLICANT: NISHIKAWA, Satomi	
; APPLICANT: OEDA, Kenji	
; TITLE OF INVENTION: Plant Promoter	
; FILE REFERENCE: 2185-0353P	
; CURRENT APPLICATION NUMBER: US/09/352,608	
; CURRENT FILING DATE: 1999-07-13	
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN	
; EARLIER FILING DATE: 1998-07-15	
; NUMBER OF SEQ ID NOS: 20	
; SOFTWARE: PatentIn Ver. 2.0	
; SEQ ID NO 2	
; LENGTH: 246	
; TYPE: DNA	
; ORGANISM: Daucus carota	
; US-09-352-608-2	
Query Match 9.5%; Score 195.6; DB 3; Length 246;	
Best Local Similarity 89.5%; Pred. No. 1.7e-32;	
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;	
QY 1805	TCTAGAATACATCTTTTCAAAATTTCAACAAACAGCTTTTAACTTTCTTTCAACGGAT 1864
Db	
QY 1	TCTAGAATATATCTTTTGAATTTTCAACAAACAGCACTTAACCTTTCTTTTAAACAGAT 60
Db	
QY 1865	GGAATCCTTTTCTAAACCTTTTAAATAAATAAATAATGCATTAATTTATCAACA 1924
Db	
QY 61	AGAATCGTTTCTTAAACCTTTTAAATTT--AAAAAATACATTAATTAATTTATCAACA 118
Db	

```
Qy 1925 CCTCAACATTCATGCTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCACA 1984
Db 119 CCTCAACATTCATGCTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCACA 178
Qy 1985 TCAATCTTACACCAACAAACCTTGAGCTTAATTTTCTACTTATTCTCAGCAATAACATTC 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAAATTTTAGCAAAACATTC 238
Qy 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 4
US-09-352-608-6
; Sequence 6, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353p
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-6

Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 2.8e-16;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1918 ATCAACACCTCAACATTCATGCTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 1977
Db 140 ATCAACACCTCAACATTCATGCTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 81
Qy 1978 CATCATCATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTTATTCTCAGCAAT 2037
Db 80 CATCATCATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAAATTTTAGCAAA 21
Qy 2038 AACATTCCTAAATATC 2052
Db 20 AACATTCCTAAAGGTC 6

RESULT 6
US-09-949-016-12776
; Sequence 12776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 12776
; LENGTH: 187169
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(187169)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

Query Match 4.0%; Score 81.4; DB 4; Length 187169;
Best Local Similarity 45.3%; Pred. No. 1.4e-07;
Matches 417; Conservative 0; Mismatches 496; Indels 8; Gaps 3;

Qy 883 ATTTCTTTAAAGTAATGTAAACCTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAA 942
Db 465 ATATATATAAAAATATAATATAATATAAATAAATAAATAAATAAATAAATAAATAAATA 524
Qy 943 CTTTTCACAGAAATTTATTTAAACGAAACATTTTATAACATGCTCTCGGGTGTCAAT 1002
Db 525 ATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 584
Qy 1003 ATAATAGGATCACTTACTCATCCATTAACACCTTGTAAACAAATTCATAGAT 1062
Db 585 TAAATATATATCTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 644
Qy 1063 AAAATATCTTACAAATGAAAAGAGGCAATGTCTCTTTGAAAAACAAATAGGTACTCCC 1122
Db 645 ATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 704
Qy 1123 TCCGTCCCTCTGAAATGTATACATATGATTTGGACACGGAGACTAAGAAAAATGTATAA 1182
Db 1182 TCCGTCCCTCTGAAATGTATACATATGATTTGGACACGGAGACTAAGAAAAATGTATAA 1182
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-7

Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 2.8e-16;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1918 ATCAACACCTCAACATTCATGCTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 1977
Db 140 ATCAACACCTCAACATTCATGCTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 81
Qy 1978 CATCATCATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTTATTCTCAGCAAT 2037
Db 80 CATCATCATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAAATTTTAGCAAA 21
Qy 2038 AACATTCCTAAATATC 2052
Db 20 AACATTCCTAAAGGTC 6

RESULT 6
US-09-949-016-12776
; Sequence 12776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 12776
; LENGTH: 187169
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(187169)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

Query Match 4.0%; Score 81.4; DB 4; Length 187169;
Best Local Similarity 45.3%; Pred. No. 1.4e-07;
Matches 417; Conservative 0; Mismatches 496; Indels 8; Gaps 3;

Qy 883 ATTTCTTTAAAGTAATGTAAACCTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAA 942
Db 465 ATATATATAAAAATATAATATAATATAAATAAATAAATAAATAAATAAATAAATAAATA 524
Qy 943 CTTTTCACAGAAATTTATTTAAACGAAACATTTTATAACATGCTCTCGGGTGTCAAT 1002
Db 525 ATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 584
Qy 1003 ATAATAGGATCACTTACTCATCCATTAACACCTTGTAAACAAATTCATAGAT 1062
Db 585 TAAATATATATCTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 644
Qy 1063 AAAATATCTTACAAATGAAAAGAGGCAATGTCTCTTTGAAAAACAAATAGGTACTCCC 1122
Db 645 ATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 704
Qy 1123 TCCGTCCCTCTGAAATGTATACATATGATTTGGACACGGAGACTAAGAAAAATGTATAA 1182
Db 1182 TCCGTCCCTCTGAAATGTATACATATGATTTGGACACGGAGACTAAGAAAAATGTATAA 1182
```


Db 705 ATAATATATATAAAATATATAATATATATATATATAAAATATATAAAATATATAATAT 764
Qy 1183 GTAATGTAGAGTAAAAAGAAAGAGAAAGTGGGTAAAGTAGCGGACCCACCAATA 1242
Db 765 ATATAAAATATAAAATATATATATATATATATATATAAAATATATAAAATATATAAA 824
Qy 1243 TATAAATTTGATAGATTTTGAAGAAAGTAGTTGAAGTAGTGGGTGGGATTTTATATTA 1302
Db 825 TATAATATATAAAATATATATATATATAAAATATATAAAATATATAAAATATATAAA 878
Qy 1303 TAAAAATTTACTATTTTGAAGAAAGTTTGAAGATGTATAGAAATTTGAGTGGGACATCA 1362
Db 879 TATAAAATATATAAAATATATATATATAAAATATATAAAATATATAAAATATATAAA 938
Qy 1363 AAGGAAGTGTATAGAAATTAATGGGACAGGAGTAAATACCTTTTATGATATATAAAT 1422
Db 939 ATAAATATATATAATAAAATATATATGGATATATAATATATATATATATATA-AAA 997
Qy 1423 TTTGTTATTTTGTATTAAGATTTATAATCTATGTTTATAATGATAATATATATTTTAA 1482
Db 998 TATAATATATAAAATATAAAATATAAAATATATAATATAAAATATATAATATAAA 1057
Qy 1483 AATAAATCTATATATAATCTGATTTAGTTCGATTCAGGCTTTTATAAATTTTACAATCTGA 1542
Db 1058 TAT-ATAATGTATAAAATATAAATATATAATATAAATATAAATATAAATATAAATATA 1116
Qy 1543 GTAATATGAATTAATCAGTTATCTGAAAGCAATATATATCTTTTGAAGACAGGCTTCGG 1602
Db 1117 ATATAAATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 1176
Qy 1603 TCAAAATGGGAAGTTCATGTCTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT 1662
Db 1177 TTATATATATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 1236
Qy 1663 TGTATTTTGTTCAGAAATTTAAATATAAATATAAATATAAATATAAATATAAATATAA 1722
Db 1237 TATAATATAAATAATATATATTTTATATAATATAAATATAAATATAAATATAAATATA 1296
Qy 1723 ATTGAGCAGCACTAGACTGTTTGAACATGTATGTCGGGTGATACATCTATGACCTTTCAA 1782
Db 1297 ATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 1356
Qy 1783 CTCAAACTAGTGAATATGCA 1803
Db 1357 TATATTTTAGTTTCTAAGACA 1377

RESULT 7
US-09-949-016-15940
; Sequence 15940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 15940
; LENGTH: 191569
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(191569)

; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15940
Query Match 4.0%; Score 81.4; DB 4; Length 191569;
Best Local Similarity 45.3%; Pred. No. 1.4e-07; Indels 8; Gaps 3;
Matches 417; Conservative 0; Mismatches 496;
Qy 883 ATTTCCTTTAAAAAGTAAATGTAAACCTGATCTATAAGCACATAGAAAATTTAGTACAGGTTAAAA 942
Db 465 ATATATATATAAAATATATAATATATAAAATATAAATATAAATATAAATATAAATATAA 524
Qy 943 CTTTTCAGAAGATTTATATATAACGAAATCAATTTTATAACATGCTCTCGGCTGTCAAT 1002
Db 525 ATAATATATAAATATATAATATAAATATAAATATAAATATAAATATAAATATAAATATA 584
Qy 1003 ATAATAGGATCACCTTACTGATCAATCCATTAACCTTGAACCAATTTCAATCAGAT 1062
Db 585 TAAATATATATATCTATTAATATAAATATAAATATAAATATAAATATAAATATAAATATA 644
Qy 1063 AAAATATCTTACAATGAAGAAAGGACATGCTCTTTGAAAAAACAATAGGTACTCCC 1122
Db 645 ATATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 704
Qy 1123 TCCGTCCCTCTGAAATGTATACATATGGATGGACACGAGACTAAGAAAAATGTATATA 1182
Db 705 ATAATATATAAATAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 764
Qy 1183 GTAATGTAGATTAAGAAAGAAAGAAAGTGGTAAAGTAGCGGACCCACCAATA 1242
Db 765 ATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 824
Qy 1243 TATAATTTGATAGATTTTAGAAAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATTA 1302
Db 825 TATATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 878
Qy 1303 TAAAAATTTACTATTTTGAAGAAAGTTTGAAGATGTATAGAAATTTGAGTGGGACATCCATA 1362
Db 879 TATAAATAATATAAATAATATAAATATAAATATAAATATAAATATAAATATAAATATA 938
Qy 1363 AAGGAAGTGTATAGAAATTAATTTGAGGACAGAGGAGTAAATACCTTTTATGATATATAA 1422
Db 939 ATAAAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 997
Qy 1423 TTTGTTATTTTGTATTAAGATTTTACAAGATTAATCTATGTTTATAATGATAAATATTTTAA 1482
Db 998 TATAATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 1057
Qy 1483 AATAATCTATATATTAATCTCGATTTAGTTCGATTCAGGCTTTTATAAATTTTACAATCTGA 1542
Db 1058 TAT-ATAATGTATAAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 1116
Qy 1543 GTAATATGAATAAATCAGTTATCTGAAAGCAATATATATCTTTTGAAGACAGGCTTCGG 1602
Db 1117 ATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 1176
Qy 1603 TCAAAATGGGAAGTTCATGTCTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT 1662
Db 1177 TTATATATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 1722
Qy 1663 TGTATTTTGTTCAGAAATTTAAATATAAATATAAATATAAATATAAATATAAATATAA 1782
Db 1237 TATAATATAAATAATATATTTTATATATATAAATATAAATATAAATATAAATATAA 1296
Qy 1723 ATTGAGCAGCACTAGACTGTTTGAACATGTATGTCGGGTGATACATCTATGACCTTTCAA 1782
Db 1297 ATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 1356
Qy 1783 CTCAAACTAGTGAATATGCA 1803
Db 1357 TATATTTTAGTTTCTAAGACA 1377

US-09-949-016-15851
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15851
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

Query Match 3.5%; Score 72.2; DB 4; Length 205044;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

Qy 1135 AAATGTATACATATGGATTGGACACGAGACTAAGAAAAATGTATAAAGTAAGTACAGT 1194
Db 200520 ATATATTTATATATATTTTAAATATAATATAAATAATATATTTATATATTTTAAAT 200579

Qy 1195 AAAAAGAAAGAGAAAGAGTGGGTAAAGTAGCGGGACCCACCAATATATAA----TTG 1250
Db 200580 ATAAATATATAAAATATATTTTATATATATTTTAAATATAAATATAAATAATATTTAT 200639

Qy 1251 ATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAAATTT 1310
Db 200640 ATATATTTTAAATATAAATATATAAATATATATTTTATATATATTTTAAATATAAATATA 200699

Qy 1311 TACTATTTTGAGAAAGTTTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAG 1370
Db 200700 AATATATTTTATATATATTTTAAATATAAATATAAATATAATTTATATATTTTAA 200759

Qy 1371 TGTATAGAAATTAATGGGACAGAGGAGTAATACCTTTTATGATATATAAATTTTGTAT 1430
Db 200760 TATAAATATATAAATATATTTTATATATATTTTAAATATAAATATAAATAATATATA 200819

Qy 1431 TTTGATTTTCATAGATTAATAATCTATGTTTATATGATATATAATTTTAAATATAATAC 1490
Db 200820 TTTTAAATATAAATATAAATAATATAATATTTTAAATATAAATATAAATAATATA 200879

Qy 1491 TATATTAATTCGATTAGTCGATTCGCCCTTTTATAATTTTACAATACCTGAGTAATG 1550
Db 200880 TATTTTAAATATAAATATAAATAATATAATATTTTAAATATAAATATAAATAATATA 200939

Qy 1551 AATAAATCAGTTATCTGAAAGCAAAATAATCTCTTTGTAAAAACAGCGTTCGGTCAA---- 1606
Db 200940 TATATTTTAAATATAAATAATATAAATAATATAATTTTAAATATAAATAATATAAATA 200999

Qy 1607 -----ATGGGAAGTTCATGTTTCAATAGTTTAAATATAAAGTAATTTTAAATTA 1661
Db 201000 TATATATTTTAAATATAAATAATATAAATAATATAATTTTAAATATAAATAATATAA 201059

Qy 1662 TTGTTATTTTGTTCAGAAATTTTAAATAAATTTTATGA 1700
Db 201060 TATATATATTTTAAATATAAATAATATAAATAATATTTTA 201098

RESULT 9

US-09-949-016-15852
; Sequence 15852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15852
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15852

Query Match 3.5%; Score 72.2; DB 4; Length 205044;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

Qy 1135 AAATGTATACATATGGATTGGACACGAGACTAAGAAAAATGTATAAAGTAAGTACAGT 1194
Db 200520 ATATATTTATATATATTTTAAATATAAATATAAATAATATTTATATATTTTAAAT 200579

Qy 1195 AAAAAGAAAGAGAAAGAGTGGGTAAAGTAGCGGGACCCACCAATATATAA----TTG 1250
Db 200580 ATAAATATATAAAATATATTTTATATATATTTTAAATATAAATATAAATAATATTTAT 200639

Qy 1251 ATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATAAATTT 1310
Db 200640 ATATATTTTAAATATAAATATAAATATATAATTTTATATATATTTTAAATATAAATATA 200699

Qy 1311 TACTATTTTGAGAAAGTTTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAG 1370
Db 200700 AATATATTTTATATATATTTTAAATATAAATATAAATATAATTTATATATTTTAA 200759

Qy 1371 TGTATAGAAATTAATGGGACAGAGGAGTAATACCTTTTATGATATATAAATTTTGTAT 1430
Db 200760 TATAAATATATAAATATATTTTATATATATTTTAAATATAAATATAAATAATATATA 200819

Qy 1431 TTTGATTTTCATAGATTAATAATCTATGTTTATATGATATATAATTTTAAATATAATAC 1490
Db 200820 TTTTAAATATAAATATAAATAATATAATATTTTAAATATAAATATAAATAATATA 200879

Qy 1491 TATATTAATTCGATTAGTCGATTCGCCCTTTTATAATTTTACAATACCTGAGTAATG 1550
Db 200880 TATTTTAAATATAAATATAAATAATATAATATTTTAAATATAAATATAAATAATATA 200939

Qy 1551 AATAAATCAGTTATCTGAAAGCAAAATAATCTCTTTGTAAAAACAGCGTTCGGTCAA---- 1606
Db 200940 TATATTTTAAATATAAATAATATAAATAATATAATTTTAAATATAAATAATATAAATA 200999

Qy 1607 -----ATGGGAAGTTCATGTTTCAATAGTTTAAATATAAAGTAATTTTAAATTA 1661
Db 201000 TATATATTTTAAATATAAATAATATAAATAATATAATTTTAAATATAAATAATATAA 201059

Qy 1662 TTGTTATTTTGTTCAGAAATTTTAAATAAATTTTATGA 1700
Db 201060 TATATATATTTTAAATATAAATAATATAAATAATATTTTA 201098

Qy	1662	TTGTTATTTT	TGTTGTTT	CAGAAATT	TAAAAA	TAAAT	TATTGA	1700
D _b	171487	TATATATATTTT	TAAATATATA	TAAATATATA	TATATA	TAAAAA	TATATTTA	171525

RESULT 14
US-08-232-463-14/C

```

; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

Query Match	3.4%	Score 69.4;	DB 1;	Length 7218;
Best Local Similarity	4.6%;	Pred. NO. 2.4e-05;		
Matches 16;	Conservative 212;	Mismatches 123;	Indels 0;	Gaps 0;

QY 1051 ATTCAATGATAAATATCTTACAATGAAGAAGCAATGTCCTTTGAAAAACAA 1110

DB	1461	AGTTAAAGAGATAGAAGAATTGGTAC	1170	ATGGTACTCCTCCGTCCTCTGAAATGTATACATATGGATTGGACACGGAGACTAAGA
QY				

[illegible][illegible][illegible]

```

RESULT 15
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEL
US-09-806-708B-22

```

Query Match	3.3%	Score 58.6;	DB 4;	Length 1141;
Best Local Similarity	13.1%	Pred. No. 2.4e-05;		
Matches 141;	Conservative	392;	Mismatches 536;	
Indels 10;	Gaps			

QY 538 TTTTGGCGGAATGCTTCTCAAAATGTTTTTATGTAAATAATGCCCATCCAAGGATA 597
|||:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

DB 1129 TTTKTKYKANNNNNNNGMGKDWNRMDATKWSATGTAAWTTNHAKRGAATCMCWYWTGTN 1070

QY 598 AGTAAAAATCCCGCTTTAAACCAAGTTTGTTAAATATATATCTTTACACTTACAAGAGGATATT 657

Db	1069	RRRCMTTYAMRTWTYTRSNANWSCATKHWTTMKYATKIRTA-WYAMWCAWRNNNMWCAT	1011
Ov	658	CGTAATACTTTTATAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGTGTAAACAGCGCTAG	717

Db 1010 NGYAKSCATNNMYYATTRWAAAYAAAKWARWAGNNMRYGAAAGNKGWCMAANATMGBMW 951

DQ
718 ACTTTGGTCACTGATAAATAGTAATTGTAGTAGTAAATAGTAGGATCTACAAATGCACAT 777

DB
950 ADTAGKMCNNNNNWTIDVRMMAM--KAKNNNNNNAYWTCYNRAATNNKMAATHMKWKTH 893

QY 778 AAAATTAGAGCTATTAAATTAAGTTACTAATAATAATAGAGAGGTTAGTTAGTAAACAGAAAGCAG 8337

Dh 892 GAHSKRTRHHHTTCRRTCGNNNNNARTVYWHHAARWMNAWTR TNNNNNNNNNNAC 833

QY 838 GTAAAAACAAGAGCTTGGCTGCTGTGTTAGTTGTTGCTGAGCTCATTTCTTTAAAGTA 897

D832 RNRTRW--WABWKHSWCNNNNNNNNNNNNNTWCHYTTANABBCYRANNNNNAARMARTC 775

D998 ATGTAAACTGATCTAAGCACATAGAAATTAGTACAGGTTTAAACTTTTACAGAATTT 957

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 18:28:54 ; Search time 1188 Seconds
(without alignments)
10274.728 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgcctacagcaca.....gcataacattcctaataatc 2052

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99.4	4.8	8056	18	US-10-473-126-386
2	93	4.5	3673778	16	US-10-312-841-1
3	91.8	4.5	8056	18	US-10-473-126-240
4	91.6	4.5	8056	18	US-10-473-126-386
5	88.6	4.3	8056	18	US-10-473-126-240
6	83	4.0	158001	17	US-10-211-179-11
7	79.4	3.9	74665	18	US-10-719-993-6854
8	75	3.7	11745	15	US-10-240-453-206
9	71	3.5	5930	15	US-10-311-455-490
10	70.6	3.4	3673778	16	US-10-312-841-1
11	70	3.4	3252	13	US-10-027-632-113786

c 12	70	3.4	3252	13	US-10-027-632-113787	Sequence 113787,
c 13	70	3.4	3252	13	US-10-027-632-113788	Sequence 113788,
c 14	70	3.4	3252	17	US-10-027-632-113786	Sequence 113786,
c 15	70	3.4	3252	17	US-10-027-632-113787	Sequence 113787,
c 16	70	3.4	3252	17	US-10-027-632-113788	Sequence 113788,
c 17	69.8	3.4	6352	17	US-10-221-613-195	Sequence 195, App
c 18	69.4	3.4	16258	17	US-10-257-166-120	Sequence 120, App
c 19	69	3.4	3673778	16	US-10-312-841-2	Sequence 2, Appli
c 20	68.6	3.3	49979	18	US-10-741-601-5746	Sequence 5746, Ap
c 21	68.6	3.3	49979	19	US-10-741-600-17905	Sequence 17905, A
c 22	68.2	3.3	15548	15	US-10-311-455-2128	Sequence 2128, Ap
c 23	68	3.3	11422	15	US-10-311-455-191	Sequence 191, App
c 24	68	3.3	11422	17	US-10-257-166-17	Sequence 17, Appl
c 25	67.8	3.3	15732	14	US-10-239-676-95	Sequence 95, Appl
c 26	67.8	3.3	15732	15	US-10-240-453-107	Sequence 107, App
c 27	67.6	3.3	5984	18	US-10-433-793-24	Sequence 1984, Ap
c 28	67.6	3.3	6050	15	US-10-311-455-1184	Sequence 1115, Ap
c 29	67.4	3.3	5748	15	US-10-311-455-1115	Sequence 1115, Ap
c 30	67.4	3.3	18154	15	US-10-311-455-228	Sequence 228, App
c 31	67.2	3.3	5807	15	US-10-311-455-1128	Sequence 1128, Ap
c 32	67.2	3.3	6816	18	US-10-723-860-7676	Sequence 7676, Ap
c 33	67	3.3	2053	13	US-10-027-632-97533	Sequence 97533, A
c 34	67	3.3	2053	17	US-10-027-632-97533	Sequence 97533, A
c 35	67	3.3	115218	19	US-10-278-698-255	Sequence 255, App
c 36	67	3.3	115218	19	US-10-278-698-769	Sequence 769, App
c 37	66.6	3.2	419	9	US-09-960-352-11234	Sequence 11234, A
c 38	66.6	3.2	6079	15	US-10-311-455-394	Sequence 394, App
c 39	66.6	3.2	6292	17	US-10-221-714A-461	Sequence 461, App
c 40	66.4	3.2	18624	15	US-10-311-455-1676	Sequence 1676, Ap
c 41	66.2	3.2	1501	18	US-10-473-126-328	Sequence 328, App
c 42	66.2	3.2	3252	13	US-10-027-632-113786	Sequence 113786,
c 43	66.2	3.2	3252	13	US-10-027-632-113787	Sequence 113787,
c 44	66.2	3.2	3252	13	US-10-027-632-113788	Sequence 113788,
c 45	66.2	3.2	3252	17	US-10-027-632-113786	Sequence 113786,

ALIGNMENTS

RESULT 1

US-10-473-126-386

; Sequence 386, Application US/10473126

; Publication No. US20040234973A1

; GENERAL INFORMATION:

; APPLICANT: Epigenomics AG

; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell

; TITLE OF INVENTION: proliferative disorders

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/473,126

; CURRENT FILING DATE: 2003-09-26

; NUMBER OF SEQ ID NOS: 1258

; SEQ ID NO 386

; LENGTH: 8056

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-473-126-386

Query Match 4.8%; Score 99.4; DB 18; Length 8056;
Best Local Similarity 44.5%; Pred. No. 1.1e-07;
Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;

Qy	548	ATGCTTCTCAAAATGTTTTTATATGTAATAATGATCCCAAGGATAAGTAAATTC	607
Db	1975	ATTAATATTAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATATAT	2034
Qy	608	CCGTTAACCAAGTTTGTAATATATATATGTTTACACTTACAGAGATATTCGTAATCTT	667
Db	2035	ATTTTAAATATTTTATTTTAAAAAATTTTATTTAAAAATATTTTAAATATATAAT	2094
Qy	668	TTAGACGACAGAGACTTAGTCAAAATCGACGCTGGTAAACAGCCTAGACTTGGTCAC	727


```
Db 1714186 ATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATAAATAT 1714127
Qy 1105 AAACAAATAGGTACTCCCTCGCTCCCTGCTGAAATGATACATATGGATGGACACGGAGA 1164
Db 1714126 TATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATAT 1714067
Qy 1165 CTAAGAAAATGTATATAAGTAATGTAGATGTAAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1223
Db 1714066 ATATATAAATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATAT 1714007
Qy 1224 GTAGCGGACCCCAATATATATATATATATATATATATATATATATATATATATATATATAT 1283
Db 1714006 TATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATAT 1713947
Qy 1284 GGGTGGGATTTTATATATATAAATATATATATATATATATATATATATATATATATATATAT 1343
Db 1713946 ATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATAT 1713887
Qy 1344 TTGAGTGGGACATCCCAATAAAGAAAGGTATAGAAATTAATGGGACAGAGGGAGTAAATA 1403
Db 1713886 ATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAAT 1713827
Qy 1404 CCTTTATGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1463
Db 1713826 TATATAAATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATA 1713767
Qy 1464 ATGATATATAAATTTTAAATATATATATATATATATATATATATATATATATATATATATAT 1519
Db 1713766 TATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATAT 1713707
Qy 1520 CTTTTATAATTTTACAATCTAGTAATATGAATATAATCAAGTTATCTGAAAGCAATATA 1579
Db 1713706 AATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATAT 1713647
Qy 1580 TATCTTTGTAAGACAGGTTCCGTCAAATGGGAGTTCATGTGTATTCATATGTTTTTAAT 1639
Db 1713646 AATATATATAAATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATAT 1713587
Qy 1640 ATAAAGTAAATTTTAAATTAATTTGTTATTTTGTGTTTTCAGAAATTTAAATTAATTTG 1699
Db 1713586 ATAAATATATAAATATATAAATATATATAAATATATATAAATATATATAAATATATATA 1713527
Qy 1700 AGCATGGGAAGTTCCAGGGCATCATTCGACGACCTAGACTGTTTGAACAATGTATCTCC 1759
Db 1713526 AATTATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAT 1713467
Qy 1760 GGTGTACATCTATGACCTTTCACTCAACTCAACTAGTAATGCAATTCAGAAATACATCTT 1819
Db 1713466 ATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATA 1713407
Qy 1820 TTCAAAATTCACAAACACACAGCTTTAACTTTTCTTCAACGGATTTGGAATCCTTTTCTAA 1879
Db 1713406 TAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATAT 1713347
Qy 1880 ACTTTTTAAATATAAATAA 1898
Db 1713346 AATATATATAAATATATA 1713328
```

RESULT 3

```
US-10-473-126-240
; Sequence 240, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 240
```

```
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-240
```

```
Query Match 4.5%; Score 91.8; DB 18; Length 8056;
Best Local Similarity 44.2%; Pred. No. 2.4e-06;
Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5;

Qy 547 AATGCTTCTCAAAATGTTTTTATATATGTAATAAATCCCATCCCAAGGATTAAGTAAATTT 606
Db 1358 ATTTTTTATTAATAAATTTATTTTTTTTATAAATAATAATATCGATTTTTTTTTTATTTT 1417
Qy 607 CCGTGTAAACAGTTCTTAATATATATATGTTTACCTTTACAAGAGGATATTCGTAATACT 666
Db 1418 TTTATTTTTTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1477
Qy 667 TTTAGACGACAAGAGACTTAGGTCAAAAATGGCGTGGTAAACAGCTAGACTTGGTCA 726
Db 1478 ATAATTAATTTAATAATTAATTTATATAAATAAATAAATAAATAAATAAATAAATAAAT 1537
Qy 727 CTGATAAATAGATAATTTGTTAGTATATATATAGTAGGATCTCAATGACATTAATAATAGA 786
Db 1538 TTAATAAATAATTTATTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1597
Qy 787 GCTATTAAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAACAGAAAAGAGGTAATAACA 846
Db 1598 TAATTTTAATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1657
Qy 847 AGAGCTTGCTGCTGTGTTTGTGTTGTTGAGCTCATTTCTTTAAAGTAATGTAACACT 906
Db 1658 AATTAATAC-GATAAATAATTTATTTTTTATAAATAAATAAATAAATAAATAAATAAATAA 1716
Qy 907 GATCTAAAGCACATAGAAATTTAGTACAGGTAAACCTTTTACAGAGATTTATATTAATAAC 966
Db 1717 TATATATTTTAAATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1776
Qy 967 GAAAAATCATTTTAAACATGCTCTCGGCTGTCTATTAATAATAGGATCACCTTACTGATCA 1026
Db 1777 AATTAAATAATTAATACGTTATATATTTTAATTAATTAATAAATAAATAAATAAATAAATA 1836
Qy 1027 TCCATTAATAAACCCTTTGTTAAACCAATTCATAGAGATAAATAATCTTCAATGAAAGAG 1086
Db 1837 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1896
Qy 1087 GACAAATGCTCTTTGAAAAAACAATAAGGTACTCCCTCCGCTCTGAAATGTATACAT 1146
Db 1897 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1956
Qy 1147 ATGGAATGGACACGGAGACTAAGAAATACTATAAAGTAATGTAGAGTAAATAAAGAGAG 1206
Db 1957 AATTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2016
Qy 1207 A-----AAGAAAGTGGTAAAGTAGCGGGACCCCAATATATATATTCATAGAT 1256
Db 2017 ATATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2076
Qy 1257 TTAGAAAAAGTAGTTGAAGTAGTGGGTGGGTGGGATTTTATATATATAAATAAATAAATAA 1316
Db 2077 ATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2136
Qy 1317 TTTGAGAAAGTTTGAATGCTATAGAAATTCAGTGGGACATCCATAAAGGAAAGTGTATA 1376
Db 2137 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2196
Qy 1377 GAATTAATGGGACAGAGGAGTAAATACCTTTATGATATATAAATAAATAAATAAATAAATAA 1436
Db 2197 ACGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2256
Qy 1437 TTCATAAGATTATAAATCTATGTTAATGATATAAATAAATAAATAAATAAATAAATAAATAA 1496
```


Db 127733 TATATAATATAATATAATAAATATAATAATATAATAATATAATAATAATAATAATAT 127792
Qy 1544 TAATATGAATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTTCGGT 1603
Db 127793 AATATAATATAATAAT 127849
Qy 1604 CAAATGGGAAGTTTCATGTGTATTTCAATAGTATTTTAAATATAAAGTAAATTTTAAATTTAAAT 1663
Db 127850 TATATATAGCATATAAATAATATATATACATATATATAAATAATATATATATATATATAT 127909
Qy 1664 GTTATTTTTGTTTCAGAAATTTTAAATAAAATTAAT 1698
Db 127910 ATTAT 127944

RESULT 7

US-10-719-993-6854
; Sequence 6854, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 53342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6854
; LENGTH: 74665
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-6854

Query Match 3.9%; Score 79.4; DB 18; Length 74665;
Best Local Similarity 44.8%; Pred. No. 0.00085;
Matches 584; Conservative 0; Mismatches 691; Indels 30; Gaps 6;

Qy 637 TTACACTTACAAGAGGATTCGTATATACATCTTTAGACGACAGACACTTAGGTCAAAAT 696
Db 54564 TTTGTCTCAAAAATAAT 54623
Qy 697 GGACGCTGTTAAACAGCCTAGACTTGGTCACTGATAAATAGATAATTTGTTAGTATAATAT 756
Db 54624 AATGTGTAATATATTAATTAAT 54683
Qy 757 AGTAGGATCTCAATGACATTAATAATTTAGAGCTATTAATTTAAGTTACTATAATAATAAGAG 816
Db 54684 ATTTATATATGTAATATAAT 54743
Qy 817 AGTTAGTAAACAGAGCAGGTAAACACAGAGCTTGCCTGCTGCTGTTAGTTGTTGT 876
Db 54744 ATTAAATAATATATTTATATATATGTAATATATATATATATATATATATATATATATATAT 54803
Qy 877 GAGCTCATTTCTTTAAAGTAATGTAACCTGATCTAAAGCACATAGAAAATTTAGTACAGG 936
Db 54804 AATATGTAATATATTAATTAAT 54863
Qy 937 TTAACACTTTTACAAGAAATTTATATTAACGAAATCAATTTTATAACATGCTCTCGGCT 996
Db 54864 ATTATATATGTAAT 54923
Qy 997 GTCATTTAATAGGATCATTCTACTGATCATCATTAACACCTTGTAAACAAATTCAT 1056
Db 54924 TAAATATATATATATATATGTAATATATATATATATATATATATATATATATATATATAT 54982
Qy 1057 TGAGATAAATAATCTTACAAATGAAGAAGCAATGCTCTTTTGAAACAAATAGGT 1116
Db 54983 TAGTTAAAT 55042
Qy 1117 ACTCCCTCCGCTCGCTCGAAATGTATACATATGATTTGGACGAGACTAAGAAAATG 1176
Db 55043 AATATGTAATATATTAATAAT 55089

Qy 1177 TATAAAGTAATCTAGAGTAAAAAGAAAGAAAGTGGTAAAGTAGCGGGGCCCA 1236
Db 55090 TATTAATAAT 55149
Qy 1237 CCAATATATAAATTTGATAGATTTTAGAAAAGTAGTGTGAAGAGTAGTGGGTGGGATTTTT 1296
Db 55150 TAATATATTAATAAT 55209
Qy 1297 ATATTTATAAATAATTTACTATTTTGGAAAAGTTTGAATTTGTATAGAAATTTGAGTGGGACAT 1356
Db 55210 ATATATTAATAAT 55263
Qy 1357 CCATAAAGGAAAGTGTATAGAAATTAAT- GGCACAGAGGGAGTATACCTTTATGATAT 1415
Db 55264 TTATATATGTAATATGTAAT 55323
Qy 1416 ATAAATTTTTTGTATTTTGTATTTTCAAGATTTATAATCTATCTGTTATAATGAATAATATA 1475
Db 55324 AATTATATATGTAATATGTAAT 55383
Qy 1476 TTTTAAATAATATCTATATATATTTCTGATTTAGTCGATTTACCGCCTTTTATATATTTTACA 1535
Db 55384 TATRTAATAATATTAATAATATATTTATATATATGATATGTAATATATATATATATATAT 55443
Qy 1536 ATACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAAATAATCTTTTGTAAAACAG 1595
Db 55444 TATATGTAAT 55503
Qy 1596 CGTTCGGTCAAAATGGGAAGTTTCATGTGTTCAATAGTTTAAATATAAAGTAAATTTTTA 1655
Db 55504 ACTACATTTTAAATATATATATTAATAATATATATATATATATATATATATATATATAT 55563
Qy 1656 AATTAAATTTGTTATTTTGTTCAGAAAATTTAAATAAATTTATGACATGGGAGTTTCCAC 1715
Db 55564 AATATATATATATATATTAATAATATATATATATATATATATATATATATATATATAT 55619
Qy 1716 GGGCATCATTTGAGCAGCAGTACAGTCTTTGAAACAATGTATGTCGGGTGACATCTATGAC 1775
Db 55620 ----ATATATTAATTAATATATATCTGATGTTAAATATATATATATATATATATATAT 55675
Qy 1776 CTTTCAACTCAAACTAGTGAATAAGTCAATTTCTAG-AATACATCTTTTCAAAATTTCAACAA 1834
Db 55676 TATACAGTATATATTTAT 55735
Qy 1835 ACACAGCTTTAACTTTTCTTTTCAACGATTTGGAATCCTTTCTAAACTTTTTHAAATAAA 1894
Db 55736 AATAATATATATCATGTAATTAATTTATATATATATATATATATATATATATATATATAT 55795
Qy 1895 AAAAAATGCAATTTTGTAAATTTTATCAACACCTTCAACATTTGATGT 1939
Db 55796 TTATATATAGTAATGTAAT 55840

RESULT 8

US-10-240-453-206/c
; Sequence 206, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06

;; PRIOR APPLICATION NUMBER: DE 10019173.8
;; PRIOR FILING DATE: 2000-04-07
;; PRIOR APPLICATION NUMBER: DE 10032529.7
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 350
;; SEQ ID NO 206
;; LENGTH: 11745

;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;; NAME/KEY: unsure
;; LOCATION: (9105)
US-10-240-453-206

Query Match 3.7%; Score 75; DB 15; Length 11745;
Best Local Similarity 42.7%; Pred. No. 0.0027;
Matches 499; Conservative 0; Mismatches 665; Indels 5; Gaps 2;

Qy	768	CAATGACATTAATAATTAGAGCTATTAAATTAAGTTACTAAATAAAGAGAGGTTAGTAAA	827
Db	4805	CGAATAAATAATAATCTATTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4746
Qy	828	CAGAAAGCAGGTAAAAACAAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	887
Db	4745	AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4686
Qy	888	TTTAAAGTAGTAAGTAAGTCTTAAGCAGCATAGAAATTTAGTACAGGTTAAACATTTT	947
Db	4685	CTCATTACACATCTACCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4626
Qy	948	ACAGAGATTTATATTAAGCAGAAATCAATTTTATACATGTCTCTCGGCTGTCATTAAAT	1007
Db	4625	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4566
Qy	1008	AGGATCACTTACTGATCATCCATTAAACCTGTGTAAAAAATAAATAAATAAATAAATAA	1067
Db	4565	ACCTTAAATTTAACTCGAATCACCTAAAA---AAAAAACAATAAATAAATAAATAAATAA	4509
Qy	1068	ATCTTACATGAAGAAGCAATGTCTCTTTGAAAAAACAATAAGTAGTCTCCCTCGGT	1127
Db	4508	ACCAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4449
Qy	1128	CCCTCTGAATGTATACATATGATGGATTGGACAGGAGACTAAGAAATAATGTATAAGTAAT	1187
Db	4448	ACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4389
Qy	1188	GTAGAGTAAAAAGAAAGAGAAAGAAAGTGGTAAAGTAGGGGAGCCCAATATATAA	1247
Db	4388	ATAACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4331
Qy	1248	TTGTATAGATTTAGAAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGG	1307
Db	4330	TTTAAATCCTAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4271
Qy	1308	ATTTACTATTTTGTAGAAAGTTTGAATGTATAGATTTAGTGGGAGCATCAATAAAGGA	1367
Db	4270	TATTTTAAATTAATTTAAACCTCTTTTATTAATAAATAAATAAATAAATAAATAAATAA	4211
Qy	1368	AAGTGTATAGAAATTAAGGAGAGGAGTATACCTTTATCATATATATAAATTTTGT	1427
Db	4210	ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4151
Qy	1428	TATTTTGAATTTCAAGATTTATAATCTGTTTATAATGATAATATAATTTTAAATAAATAA	1487
Db	4150	TTCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4091
Qy	1488	TACTATATTAATTTCTGATTTAGTACCGCTTTTATATATTTTAAATAAATAAATAAATAA	1547
Db	4090	TTTTCTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	4031

Qy	1548	ATGAATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAAAACAGCGTTCGGTCAAA	1607
Db	4030	AAAAAACAACAATAATTTACAAAAACAATAAATAAATAAATAAATAAATAAATAAATAA	3971
Qy	1608	TGGGAAGTTCATGTGTATTCATATGTTTAAATATAAAGTAATTTTAAATAAATAAATAA	1667
Db	3970	AAACAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3911
Qy	1668	TTTTTGTTCAGAAATTTAAATAAATAATTTTGAGCATGGGAAGTTTCACGGGCATCATTTGA	1727
Db	3910	AAACTTTTCCATTTTAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3851
Qy	1728	GCAGCACTAGACTGTTTGAACAATGTATGTCGGGTGTACATCTATGACCTTTTCAACTCAA	1787
Db	3850	TAAAAATTAATAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3791
Qy	1788	ACTAGTGAATAATGCATTTCTAGATATACATCTTTTCAAAATTTCAACAACAACAGCTTTAAC	1847
Db	3790	AACATAAACAATAATTCATCAACCAAAATAAATAAATAAATAAATAAATAAATAAATAA	3731
Qy	1848	TTTTCTTTCAACGATTTGGAATCCTTTTCTAAACCTTTTAAATAAATAAATAAATAAATAA	1907
Db	3730	TTAAAAAATAAATAATTTCAAAATAATCACCTAAATAAATAAATAAATAAATAAATAAATAA	3671
Qy	1908	TGTAATATTTATCAACACCTCAACATTTGA	1936
Db	3670	TTAAATAATAAATACTACTAACAAATTTAA	3642

RESULT 9

US-10-311-455-490/c

; Sequence 490, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311.455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 490

; LENGTH: 5930

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-490

Query Match 3.5%; Score 71; DB 15; Length 5930;
Best Local Similarity 46.4%; Pred. No. 0.011;
Matches 371; Conservative 0; Mismatches 420; Indels 8; Gaps 4;

Qy	897	AATGTAACTGAATCTAAGCAGCATAGAAATTTAGTACAGGTAAACCTTTTACAGAAATT	956
Db	5154	ARTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	5095
Qy	957	TATATTAAACCAAAATCATTTTATAACATGCTCTCGGCTGTCTATTATATATAGGATCAC	1016
Db	5094	ATTTTAAAAAATTTTACTCATCAAACTTATCTAAAAACCTCAAACTTCCCAATTTATAT	5035
Qy	1017	TTACTGATCATCTTAAACCTTTGTTAAACAATAATTTCAATGAGATATAATATCTTACAA	1076
Db	5034	CTAATTTTATAATTTTCAAAACAATAATTTAAATAAATAAATAAATAAATAAATAAATAA	4975


```
US-10-027-632-113787/c
; Sequence 113787, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 113787
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113787

Query Match      3.4%; Score 70; DB 13; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.014;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

QY 621 TTGTTAATATATATGTTTACACTTACAGAGGATATTCGTAATACATTTTGGAGCACAAGA 680
DB 2314 TTAACTAAACATTTTATACCTTATTAATGAAGTTAATGATGTTTATGTTTACTGTAATATA 2255
QY 681 GACTTTAGGTCAAAATGGAGCGCTGGTAAACAGCCTAGACTTGCTACTGATATAATAGATA 740
DB 2254 TATTTAACTGCTATGTAACATATGAATGAAGTTAATATATATATTTAAATTTAATATAATAC 2195
QY 741 ATTGTTAGTATATATAGTAGGATCTACATGACATTAATTTAGAGCTATTAATTAAGT 800
DB 2194 ATTATAATATAAATTTTATAACAATATATCATATATACATATTCATATTTATATATTTAATA 2135
QY 801 TACTA-----ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAACAAAG 848
DB 2134 TATAATACCATAATACTTATATTTAAAGCTAATTTAATATATATACATTAATCTAATA 2075
QY 849 AGCTTGCTGCTGTTGTTAGTTGTTGTGAGCTCATTTCTTTTAAAAAGTAATGTAACCTGA 908
DB 2074 ATCTTTTACTTATTTGTTACTTACATTTAAATTAATATATATTTAACTTTATTA 2015
QY 909 TCTAAGCACAATAGAAATTTAGTACAGGTTA-AACTTTTACAAGAAATTTATATTAACG 967
DB 2014 TCTGGTTAACTATTATATATAGTTAAACCATATTTAAAAATATATAAATATAAATAACA 1955
QY 968 AAAATCAATTTTATAACATGCTCTCGGCTGCTCATTTAATAAGGATCACTTCTGATCAT 1027
DB 1954 TGATATATATATATATAAATTTGCAATTTTATTTATTTATATATATATATATTTAATAT 1895
QY 1028 CCATTTAAACCTTTGTTAAACAAATTCATAGATATAATCTTCAATGAAAGAAAGG 1087
DB 1894 ATAATATTTATATATATATTTTAAATAACAATATTTATATATATATATATTTCAATAAAC 1835
QY 1088 ACAATGCTCTTTGAAACAAACAAATAGGTAAGTCTCCCTCCGCTCGAATGTATACATA 1147
DB 1834 AATATTTATATATATATATTTTAAATAACAATATATATATATATTTAATATACAATATT 1775
QY 1148 TGGATTGGACACGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGA 1207

1774 TATATATTATATATATTTTAAATATACAAATTTTATATATATATATATATATTTAAATATACAATA 1715
1208 AGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATATTAATTCATAGATTTAGAAAAAGTA 1267
1714 TTTATATATATATATATTTTAAATATACAAATATTTATATATATATATATATTTTAAATATACA 1655
1268 GTTGAAGTAGTGGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTCAGAAAGT 1327
1654 TATTATATATATATATATTTTAAATATACAAATTTTATATATATATATATATTTTAAATATATAATA 1595
1328 TTTGAAATCTATAGAAATTCAGTGGGACATCCATATAAAGGAAAGCTGTATAGAAATTAATGG 1387
1594 TTTATATATATATATATTTGATGTCATACACATAATTAATATATATATATTTTCAATATA 1535
1388 GACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGTATTTTCAATAGATT 1447
1534 TTAATATATGAATACATCTATTAAATATATGAATATATGCTATCTTATATATATTAATATAAGA 1475
1448 ATAAATCTATGTTATATATGATA-ATATAATTTTAAAAAATAACTATATATTAATCTGATT 1506
1474 ATTATATATATATATTTGTTATATATATATATATATTTAATTAATATATCATTTGTA 1415
1507 AGTCGATTTACCGCTTTTATATAATTTTACAAATCTGAGTAATATGAATAAATCACTTATCT 1566
1414 TATTATGTAATTCGATTTTATATATATTAATTAATGTAATGCAATATATATATAATA 1355
1567 GAAAACAAATATATATCTTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTTCATGTGATT 1626
1354 TTATATAATATATATATATATTTTTCATATATTTTATATATATTTTAAATATGTTTAACTATA 1295
1627 CAATAGTTTAAATATAAAGTAATTTTAAATTAATTTGTTATTTTGTTCAGAAATTTA 1686
1294 TATAATAGTAGCATATTAATAGTTAATTTACTATATATTAATTTAATTTAATATAAGTTAA 1235
1687 AAATAAAT 1694
1234 CAGTAAGT 1227

RESULT 13
US-10-027-632-113788/c
; Sequence 113788, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 113788
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113788
```

Query Match		3.4%;	Score 70;	DB 13;	Length 3252;
Best Local Similarity		43.6%;	Pred. No. 0.014;		
Matches 474;		Conservative 0;	Mismatches 600;	Indels 14;	Gaps 3;
Qy	621	TTGTTAATATATATGTTTACACCTTACAGAGGATATTCGTAATACCTTTTACGACGACAAGA	680		
Db	2314	TTAATACCAAAATTTTATACCTTATATTAATGAGTTTAAATGTTTAACTGTAATATA	2255		
Qy	681	GACTTAGGTCAAAAATGGACGCTGGTAAACAGCCCTAGACTTGGTCACCTGATAAATAGATA	740		
Db	2254	TATTTAACTGCTGATGTAACATATGAATAGTTAAATATATTAATTAATATAATAC	2195		
Qy	741	ATTGTTAGTAAATATATAGTAGGATCTCAATGACATTTAAATATGAGCTATTAATTAAGT	800		
Db	2194	ATTATAATATAAATTTTAAACAATATATCAATATACATAATTTATATATATTTAATA	2135		
Qy	801	TACTA-----ATAAATAAGAGAGGTAGTAAACAAGAGGAGGTAACAAACAG	848		
Db	2134	TATAATACCAATAAATCTTATATTAAGCTAATTTAATATATATACATTAATCTAATA	2075		
Qy	849	AGCTTGCTGCTGCTGTTAGTTGTTGAGCTCAATTTCTTTAAAGTAATGTAACCTGA	908		
Db	2074	ATCTTTATACCTTATGTTAACTTACATTAATTAATTAATATATTAATTAATTAATA	2015		
Qy	909	TCTAAGCACATAGAAATTTAGTACAGGTTA-AAACCTTTTACAGAAATTTATATTAACG	967		
Db	2014	TCTGGTTAACTTATATATAGTTTAAACCATATATAAATATAAATATAAATAAATACA	1955		
Qy	968	AAATCATTTTATACATGCTCTCGGCTGCATTAATAGGATCACTTACTGATCAT	1027		
Db	1954	TGATATATATATATATTAATTTGCAATTTTATATATATATATATATATATATTAAT	1895		
Qy	1028	CCATTTAAACCTTCTTAAACAAAATTCATGAGATATAAATCTTACAAATGAAAGAGG	1087		
Db	1894	ATATATTTATATATATATATTTAATTAACAATTTATATATATATATATTTCAATAAC	1835		
Qy	1088	ACAATGCTCTTTGAAACAAAATAGAGTACTCCCTCGCTCCCTGAAATGTATACATA	1147		
Db	1834	AAATTTTATATATATATATTTAATTAACAATATATATATATATATTTAATATACAAT	1775		
Qy	1148	TGGATTGGACGAGAGCTAAGAAATGTATTAAGTAATGTAGAGTAAAGAAAGAGA	1207		
Db	1774	TATATATATATATATTTAATATACAAATTTTATATATATATATATATTTAATATACA	1715		
Qy	1208	AAGAAAAGTGGTAAAGTAGGCGGACCCACCAATATATATATTTAGATAGTTAGAAAAGTA	1267		
Db	1714	TTTATATATATATATATTTAATATACAAATTTTATATATATATATATTTAATATACAA	1655		
Qy	1268	GTTGAAAGTAGTGGGTGGGATTTTATATATATAAATTTTACTATTTTGGAGAAAGT	1327		
Db	1654	TATTTATATATATATATTTAATATACAAATTTTATATATATATATATTTAATATATAA	1595		
Qy	1328	TTTGAATGTATAGAAATGAGTGGGACATCCATAAAAGGAAGGTATAGAAATTAATGG	1387		
Db	1594	TTTATATATATATAAATGGATTGTATCATACATAAATTTATATATCAATATATTTCAATA	1535		
Qy	1388	GACAGGAGGATACCTTTATCATATATAAATTTTGTATTTTGTATTTTCAAGATT	1447		
Db	1534	TTAATATAGAACTATTAATATATAGAAATATATGCTATTTCTTATATTAATATATAA	1475		
Qy	1448	ATAAATCTATGTTATAATGATA-ATATAAATTTTAAAAATATACTATATTAATTTCTGATT	1506		
Db	1474	ATTATATATATATATGATATATATATATATATATATATATAGATTTAATAATCATTTGTA	1415		
Qy	1507	AGTCGATTTACCGCTTTTATTAATTTTCAATACCTGAGTATATAGAAATTAATCACTTATCT	1566		
Db	1414	TATTTATGTAATTTGATTTATGTTATATAAATGTAATGTAATGCAAAATTTATATAAATA	1355		
Qy	1567	GAAAGCAATAATATCTTTGTAAACAGCGCTCGGTCAAATGGGAGTTTCATGTGATT	1626		
Db	1354	TTATATATATATATATATATATTTCAATATATTTATATATTTTAAATGTTGTTAACTATA	1295		
Qy	1627	CAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTATTTTTCGTTTCAGAAATTTA	1686		

RESULT 14
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113786
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113786

Query Match		3.4%;	Score 70;	DB 17;	Length 3252;
Best Local Similarity		43.6%;	Pred. No. 0.014;		
Matches 474;		Conservative 0;	Mismatches 600;	Indels 14;	Gaps 3;
Qy	621	TTGTTAATATATATGTTTACACCTTACAGAGGATATTCGTAATACCTTTTACGACGACAAGA	680		
Db	2314	TTAATACCAAAATTTTATACCTTATATTAATGAGTTTAAATGTTTAACTGTAATATA	2255		
Qy	681	GACTTAGGTCAAAAATGGACGCTGGTAAACAGCCCTAGACTTGGTCACCTGATAAATAGATA	740		
Db	2254	TATTTAACTGCTGATGTAACATATGAATAGTTAAATATATTAATTAATATAATAC	2195		
Qy	741	ATTGTTAGTAAATATATAGTAGGATCTCAATGACATTTAAATATGAGCTATTAATTAAGT	800		
Db	2194	ATTATAATATAAATTTTAAACAATATATCAATATACATAATTTATATATATTTAATA	2135		
Qy	801	TACTA-----ATAAATAAGAGAGGTAGTAAACAAGAGGAGGTAACAAACAG	848		
Db	2134	TATAATACCAATAAATCTTATATTAAGCTAATTTAATATATATACATTAATCTAATA	2075		
Qy	849	AGCTTGCTGCTGCTGTTAGTTGTTGAGCTCAATTTCTTTAAAGTAATGTAACCTGA	908		
Db	2074	ATCTTTATACCTTATGTTAACTTACATTAATTAATTAATATATTAATTAATTAATA	2015		
Qy	909	TCTAAGCACATAGAAATTTAGTACAGGTTA-AAACCTTTTACAGAAATTTATATTAACG	967		
Db	2014	TCTGGTTAACTTATATATAGTTTAAACCATATATAAATATAAATATAAATAAATACA	1955		
Qy	968	AAATCATTTTATACATGCTCTCGGCTGCATTAATAGGATCACTTACTGATCAT	1027		
Db	1954	TGATATATATATATATTAATTTGCAATTTTATATATATATATATATATATTAAT	1895		

Qy	1507	AGTCGATTACGGCCCTTTTATAAATTTTACAATCTAGAGTAATATGAATAAATCAGTTATCT	1566
Db	1414	TATTATGTAATCGATTTTATTTATGTTTATATAAATAGTCAAAATTTATATAATAATA	1355
Qy	1567	GAAAGCAGATAATATCTTTTGTAAGAAACAGCGTTCCGGTCAAAATCGGAAGTTTCATGTCGTAAT	1626
Db	1354	TTATATAATATATAATAATAATATATTTTCATATATTTTAAATATGTTTAACTATA	1295
Qy	1627	CAATAGTGTTTAAATATAAAAGTAAATTTTAAATTTGTTATTTTGTTCAGAAATTTA	1686
Db	1294	TATAATAGTTAGCATATTAATAAGTTAAATTTACTATATTAAATTTAATAATAAGTTAA	1235
Qy	1687	AAATAAAT	1694
Db	1234	CAGTAAGT	1227

Search completed: March 14, 2005, 23:09:30
Job time : 1198 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 15:29:49 ; Search time 6848 Seconds
(without alignments)
11405.941 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcctacagcaca.....gcataacattcttaataatc 2052

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	97.8	4.8	1101	9	CNS000EVL
C 2	97.2	4.7	1101	9	CNS0039G
C 3	95.6	4.7	1542	9	AG386981 Mus muscu
C 4	95	4.6	1101	9	CNS0039G
C 5	94.6	4.6	1758	9	CU509408 SAIL_811
C 6	93.6	4.6	1896	9	CG753083
C 7	93	4.5	1202	8	CC262481 CH261-167
C 8	92.4	4.5	1101	9	CNS0021J
C 9	90.8	4.4	1101	9	CNS000EVL
C 10	90.4	4.4	1608	9	CU118721 ISB1-72J8
C 11	88.8	4.3	1067	6	CD386564
C 12	88.8	4.3	1268	9	AG347098 Mus muscu
C 13	88.8	4.3	1392	9	CG757503 P052-4-CO
C 14	88	4.3	1277	8	CC253231 CH261-180
C 15	88	4.3	1696	9	AG346840 Mus muscu
C 16	87.6	4.3	1275	9	CU033318 CH216-36F
C 17	87.2	4.2	887	9	AG526041 Mus muscu
C 18	87	4.2	1275	9	CU033318 CH216-36F
C 19	87	4.2	2087	9	AG333887 Mus muscu
C 20	86.8	4.2	1348	9	CG749499 P043-4-AO
C 21	85.6	4.2	1350	9	CU019486 CH216-5G1
C 22	85.4	4.2	1780	9	AG320553 Mus muscu
C 23	85.2	4.2	1320	9	CU103881 ISB1-42C8
C 24	85	4.1	1539	9	AG340947 Mus muscu

C 25	84.8	4.1	1254	9	AG349719
C 26	84.6	4.1	987	9	CNS014PQ
C 27	84.6	4.1	1364	9	CG757970 P053-2-CO
C 28	84	4.1	1101	9	CNS017KE
C 29	83.8	4.1	822	7	CK416977 AUF IpInt
C 30	83.6	4.1	1489	9	AG350139 Mus muscu
C 31	83.4	4.1	1101	9	CNS00FMC
C 32	83	4.0	1211	9	AG349657 Mus muscu
C 33	83	4.0	1592	9	CG750135 P044-3-D0
C 34	82.8	4.0	1210	9	CG749728 P044-1-CO
C 35	82.6	4.0	1506	9	AG278469 Mus muscu
C 36	81.6	4.0	1092	9	CNS020K7
C 37	81.6	4.0	1101	9	CNS00EO7
C 38	81.6	4.0	1355	9	AG346348
C 39	81.6	4.0	1745	9	AG338221 Mus muscu
C 40	81.4	4.0	812	8	BH178455 O11_J_02-
C 41	81.4	4.0	812	9	CNS07KRM
C 42	81.4	4.0	1050	9	CNS013NS
C 43	81.4	4.0	1391	9	CG754863 P050-2-G0
C 44	81.2	4.0	1094	9	CNS012F2
C 45	81.2	4.0	1715	9	AG288305 Mus muscu

ALIGNMENTS

RESULT 1
CNS000EVL/c 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BAC29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL069706
VERSION AL069706.1 GI:4949849

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1. .1101

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/notes="end : T7"

ORIGIN

Query Match 4.8%; Score 97.8; DB 9; Length 1101;
Best Local Similarity 34.8%; Pred. No. 8.7e-08;


```

AUTHORS      Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE        BAC end Sequences of Library MSMg01
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 1542)
AUTHORS      Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE        Direct Submission
JOURNAL      Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
              1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
              (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
              Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT      Clones are derived from the mouse BAC library MSMg01. For BAC
              library availability, please contact Kuniya Abe (abe@crc.riken.jp).
              Tsukuba Institute of Physical and Chemical Research (RIKEN) 3-1-1
              Koyadai, Tsukuba, 305-0074 Japan
              phone: 81-298-36-9189, fax: 81-298-36-9199
              e-mail: abe@crc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY      : pBACE3.6
Vector       : EcoRI
R.Site 1    : EcoRI
R.Site 2    : EcoRI
FEATURES     Location/Qualifiers
              1..1542
               /organism="Mus musculus molossinus"
               /mol_type="genomic DNA"
               /sub_species="molossinus"
               /db_xref="taxon:57486"
               /clone="MSMg01-201G10.TJ"
               /sex="male"
               /tissue types="mixture of kidney and spleen"
               /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match      4.7%; Score 95.6; DB 9; Length 1542;
Best Local Similarity 44.3%; Pred. No. 2.2e-07;
Matches 420; Conservative 0; Mismatches 523; Indels 6; Gaps 1;
QY 750 ATAATAGTAGGATCTACATGACATTAATAAATAGAGCTATTAAATTAAGTTACTATAA 809
DB 1103 ATAAGAAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1044
QY 810 ATAAGAGAGCTTAGTAAACAGAAAGCAGGTAAGGTAAGGCTGCTGCTGCTGCTGCTAG 869
DB 1043 AATAATAAANAATAAANAATAAANAATAAANAATAAANAATAAATAAANAATAA 984
QY 870 TTGTTGAGCTCATTTCTTTTAAAGTAATGTAAACTGATCTAAAGCACATAGAAATTTA 929
DB 983 TATAAATAATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 924
QY 930 GTACAGGTTAAACTTTTACAGAAATTTATTAATAACGAAATCAATTTTATACATCTCT 989
DB 923 ATAATAATAAATAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 864
QY 990 CTCGGCTGTCATATATAGGATCACTTACTGATCATCCATTAA-----AACCTTGTT 1043
DB 863 TTAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 804
QY 1044 AAAACAAATTCATAGATATAATATCTTCAATGAAGAAAGACAACTCTCTTTTGAA 1103
DB 803 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 744
QY 1104 AAAACAAATAGTACTCCCTCGCTCCCTCTGAAATGTATACATATGATGGATGGACGGAG 1163
DB 743 TAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 684
QY 1164 ACTAAGAAATATGATAAAGTAATGTAGATGATAAAGAAAGAAAGAAAGTGGGTAAA 1223
DB 683 AATAAATAAATAATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 624
QY 1224 GTAGCGGAGCCCAATATATAATTTGATAGATTTAGAAAAAGTAGTTGAAAGTAGTGGGT 1283

```

623 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 564

1284 GGCTGGGATTTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1343

563 TATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 504

1344 TTGAGTGGGACATCCATAAAGGAAAGTGTATAGATTAATTAATGGGACAGAGGAGTAA 1403

503 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 444

1404 CCTTATGATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1463

443 ATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 384

1464 ATGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1523

383 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324

1524 TATAAATTTTACAACTAGTATATGAATAAATCAGTATCTGAAAGCAATAATATATC 1583

323 ATATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 264

1584 TTTGTAACACAGCGTTTCGGTCAATCGGAAGTTTCATGTGATTCATAGTTTAAATATA 1643

263 AATAATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 204

1644 AAGTAATTTTAAATTAATTTGTTATTTTCTGTTTTCAGAAATTTAAAAATAA 1692

203 AATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 155

RESULT 4

CNS0039G 1101 bp DNA linear GSS 03-JUN-1999

LOCUS Drosophila melanogaster genome survey sequence IET3 end of BAC #

DEFINITION BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL063921

VERSION AL063921.1 GI:4941778

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR08K10"

FEATURES

source

ORIGIN		/clone lib="RPCI-98" /notes="end : TET3"	
Query Match		4.6%; Score 95; DB 9; Length 1101;	
Best Local Similarity		20.2%; Pred. No. 2.8e-07;	
Matches 173; Conservative		325; Mismatches 348; Indels 9; Gaps 2;	
Qy	1200	GAAAGAGAAAGAAAAGTGGGTAAGTAGCGGGACCCACCAATATATAATTTGATGATTTA 1259	
Db	189	GTWAKGGGAAAATARRGGGGGAARGGGGGGGRGGGAARAGRARRGRRAAARAAA 248	
Qy	1260	GAAAAGTAGTTGAAAGTAGTGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTT 1319	
Db	249	AARGRRARRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRG 308	
Qy	1320	GAGAAAGTTTGAATGTATAGAAATGTAGTGGGACATCCATAAAAGGAAGTGTATAGAA 1379	
Db	309	GARGRRRRRRRRRAAAGGRRRRRRRAGGRRRRRRRRRRRRRRRRRRRRRRRRRR 368	
Qy	1380	TTAAATGGGACAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGATTTTC 1439	
Db	369	RRGAGGGGRRGGGGGGGGGGMATATAWAWWWTTTTTTTTTAAWAAWAAATAA 428	
Qy	1440	ATAAGATTATAAATCTATGTATAATGATAATAATTTTAAATAATACTATATTAAT 1499	
Db	429	TTWAAWAAWAAATTTWAAWAAWAAWATAWTTTATWAAWAAWAAWAAWAAWAAW 488	
Qy	1500	TCTGATTAGTCGATTACCGCTTTTATAATTTTACATAGTGAATATGAATAATCA 1559	
Db	489	TTTTTTTATWATAWTTTWWTTTAAWAAWAAWAAWAAWAAWAAWAAWATAATTTWT 548	
Qy	1560	GTTATCTGAAAGCAATATATCTTTGTATAACAGCGTTCGTCAAATCGGAAGTTTCAT 1619	
Db	549	WWTYTTTAAWATAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 608	
Qy	1620	GTGATTCTCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAG 1679	
Db	609	HWYHTYAMWHMTWHTWAWWHHTTYTAAYYYYYTCMY-----YHYMHWHHAHAA 662	
Qy	1680	AAATTTAAATAAATATTGAGCAGTGGGAAGTTTCAGGGCATCATTTGAGCAGCAGCTAGAC 1739	
Db	663	AWTHTTWTWTHAYHWATYHYYYMYCAMCMCTHTCHCYYYHHYTAHTHTTHWYAHY 722	
Qy	1740	TGTTTGAAACAATGATGTCGGGTGACATCTATGACCTTCAACTCAACTAGTGAATAA 1799	
Db	723	YMYYWYAYWYMYCTACTYHHHHHHYHAYHTTWYAWAHAMWMMHHAHAAAAWAAWA 782	
Qy	1800	T--GCATTCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTC 1856	
Db	783	TTHHYHTTTHYMHHTYMYHYMYTCCYMCYTHCWHYHTYATCTWTHHHMTWTHWYH 842	
Qy	1857	AACGGATGTGAATCCCTTTTCTAACTTTTAAATAAAAAAATGCAATTTGTAATATT 1916	
Db	843	TWHHHTTTHWAWHHTTWCWWNHAHTTWATHCWCMTWMMHWHMMHMMHMMHMMHMMH 902	
Qy	1917	TATCAACACCTCAACATTGATGTAGGTACTATAAATAGGTCTCTTGTCCTCTACTA 1976	
Db	903	THMCMCHHHKCTCHHHHTTCHWMMHMMHMMHMMHMMHMMHMMHMMHMMHMM 962	
Qy	1977	TCATCATCAATCTTACACCAACCACTTGAGCTTAAATTTTCTACTTATTTCTCAGCAA 2036	
Db	963	CHYHMMHMYCCHYCTCTHTATTHYHMYCTCYHYCTWHTTWYAWWATAHAMTTAT 1022	
Qy	2037	TAACATCTTAATAT 2051	
Db	1023	WWWMMHWAHWATWW 1037	
RESULT 5			
CL509408			
LOCUS			
DEFINITION SAIL_811_H11.v3 SAIL Collection Arabidopsis thaliana genomic clone			

ACCESSION		SAIL_811_H11.v3, genomic survey sequence.	
VERSION		CL509408	
KEYWORDS		GSS.	
SOURCE		Arabidopsis thaliana (chale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS		Sessions A., Burke E., Preeating, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho P., Backaden, J., Ko C., Clarke, J. D., Corton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mittel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.	
TITLE		A high-throughput Arabidopsis reverse genetics system	
JOURNAL		Plant Cell 14 (12), 2985-2994 (2002)	
MEDLINE		22356987	
PUBMED		12488722	
COMMENT		Contact: Sessions A Applied Trait Genetics 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA Email: allen.sessions@syngenta.com ABRC Stock Number CS838276; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC). Sequences represent a pool of amplified genomic regions and not single contiguous sequences. Class: TDNA tagged. Location/Qualifiers 1..1758 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /ecotype="Columbia" /db_xref="taxon:3702" /clone="SAIL-811_H11.v3" /clone_lib="SAIL Collection" /note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"	
FEATURES		source	
ORIGIN		Query Match 4.6%; Score 94.6; DB 9; Length 1758; Best Local Similarity 34.2%; Pred. No. 3.3e-07; Matches 502; Conservative 0; Mismatches 957; Indels 7; Gaps 5; Qy 495 AAATTGACAAACCCGACACAACTCTATCCAGTACTATACATCTTTGGCGGAATGCTTC 554 Db ANANNAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 127 Qy 555 TCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAAGTAAAAATCCCGTTTA 614 Db NNAA 187 Qy 615 ACCAGTTGTTAATATATATGTTTACATCTTACAGAGGATATTCGTAATACTTTTAGACG 674 Db AANN 247 Qy 675 ACAAGAGCTTAGGTCAAAATGGACGCTGCTAAACAGCCTAGACTTGGTCACTGATAAA 734 Db ANAAANNN 307 Qy 735 TAGATAATTTGTTAGTATATATATAGTAGGATCTCAAAATGACATTTAAAAATTAGAGCTAT 794 Db NNNAAATATATATT 367 Qy 795 TTAAGTTACTATAAATAAGAGAGGTTAGTAACAGAAAGCAGGTAAGGTAAGGAGCTTG 854 Db TTTTATTTTANTAAAAANAAAAAANAAAAAANAAAAAANNNNNNNNNNNNNNNNNNN 427 Qy 855 CTGCTGTGTGTTTAGTTGTTGAGCTCATTTCTTTAAAGCTAATGTAAAGTAACTGATCTAAA 914 Db 428 NNNNAANNNNAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 487	

SAIL_811_H11.v3, genomic survey sequence.
CL509408
CL509408.1 GI:46006728
GSS
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1758)
Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B., Mittel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)
22356987
12468722
Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number C8936276; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not single contiguous sequences.
Class: TDNA tagged.
Location/Qualifiers
1..1758
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_811_H11.v3"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN

Query Match 4.6%; Score 94.6; DB 9; Length 1758;
Best Local Similarity 34.2%; Pred. No. 3.3e-07;
Matches 502; Conservative 0; Mismatches 957; Indels 7; Gaps 5;
QY 495 AAATTGACAAACCCGACACACTCTATCCAGTACTATATCTTTTGCCGATCTTC 554
Db 68 ANANNAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNATTTTTT 127
QY 555 TCAAAATGTTTTTATATGTAATAATGCGCATCAAGGATAAGTAAATTCCTGTTA 614
Db 128 NNAAANNNNNN 187
QY 615 ACCAGTTTGTATATATATGTTTACACTTACAAGAGGATATTCGTAATCTTTAGACG 674
Db 188 AANNNNNNNNNNNAANN 247
QY 675 ACAAGAGCTTAGTCAAAAATGGAGCGTGGTAAACAGCCTAGACTTGTCTACTGATAA 734
Db 248 ANAAAAANNAANNN 307
QY 735 TAGATAATTTGTAGTATAATATAGTAGGATCTACAAATGACATTAAATTTAGACTATA 794
Db 308 NNAAATATATATT 367
QY 795 TTAAGTTACTAATAATAAGAGAGGTTAGTAAACAGAAAGAGGTAAAAACAGAGCTTG 854
Db 368 TTTTATTANTAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNNNNNNNNN 427
QY 855 CTGCTGTGTTTGTGTTGAGCTCATTTTAAAGTAATGTAAGTAACTGATCTAA 914
Db 428 NNNAANNNNAAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNN 487

```
Qy 915 GCACATAGAAATTTAGTACAGGTTAAACCTTTTCAAGAAATTTATATATAACGAAATCA 974
Db 488 NNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNT 547
Qy 975 TTTTATAACATGCTCTCGGTGTCATATATAATAGGATCCTTACTGATCATCCATTAA 1034
Db 548 AANAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 607
Qy 1035 AACCTTGTGTTAAACAAATTCATAGAGATAAAATCTTCAATGAGAAAGGACATGT 1094
Db 608 NNNNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 667
Qy 1095 CTCTTTGAAACAAATAGGTACTCCCTCGCTCCCTCTGAAATGTATACATATGGATTG 1154
Db 668 ANANNNNANNNNAANNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 727
Qy 1155 GACACGGAGACTAAGAAAAATGTAAGTAATCTGACAGTAAAGAAAGAGAGAGAAA 1214
Db 728 AAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 787
Qy 1215 GTGGGTAAAGTAGCGGACCCACCAATATATATTAATTTGATAGATTTAGAAAAAGTAGTTGAAA 1274
Db 788 NAANNAANNAANNNNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 847
Qy 1275 GTAGTGGGTGGGTGGGATTTTATATATATATAATTAATTTACTATTTTGAGAAAGTTTGAAA 1334
Db 848 AATATATAAAATATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAA 907
Qy 1335 TGATAGAAATTTGAGTGGGACATCCATAAAGAGAAAGTGTATAGAAATTAATGGGACAGAG 1394
Db 908 TATTTNNAAATAATAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 967
Qy 1395 GGAGTAAATACCTTTATGATATATAAATTTTGGTTATT-TTGATTTCAATAGATTTATAAAT 1453
Db 968 TTTTATAATAAATTAATATATATATATTTTAAATTTATTTAAATTTATTAATAATAAT 1027
Qy 1454 CTATGT-TATAAGATATAATAATTTTAAATAATACTATATAAAT-CTGATTAAGTC 1510
Db 1028 TTANNNTAANAATAATAAANAATAAATAATAAATAATAAATAATAAATAATAAATAATA 1087
Qy 1511 GATTAACCGCTTTTATATATTTTCAATCTCAGTAATATGAATAAATCAGTTATCTGAAA 1570
Db 1088 ATANATTAATAANNTTTAANTATAAATAANTTTTNTTATAAATAAATANNNTTTTAAATANA 1147
Qy 1571 AGCAATAATATCTTTGTAAAAACAGCGTTCCGTCAAATGGCAAGTTTCATGTTATTCAT 1630
Db 1148 TAATAATAAANAATAAANAANNTAANTTAANATATATAAATNTTAAATAATTAATN 1207
Qy 1631 AGTTTTTAATAATAAAGTA--AATTTTAAATTAATTTGTTATTTTGTTCAGAAATTTAAA 1688
Db 1208 AATAATATNTAANTATTAATAANTNTATATAAATAATATAAATAAATAAATAATTTNT 1267
Qy 1689 ATAAATTTAGAGCATGGGAAGTTTCAGGGGATCATTTGAGCAGCAGCTAGACTGTTGAAC 1748
Db 1268 ATTAANTATTAATTAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1327
Qy 1749 AATGTATGTCGGGTGATCATCTGACCTTCAACTCAAATAGTGAATAATGATCTCTA 1808
Db 1328 TATANNNTATNTTATA-TTNTATAAANTNTAANTTAANNAATAANNTATATATATATTTNNA 1386
Qy 1809 GAATPACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTTTCAACGGATTGGAA 1868
Db 1387 AANTATNTTATNTATNTATTTTNTTATATATTTATATAAATAATATATTAATTAATAA 1446
Qy 1869 TCCTTTTCTAAACCTTTTAAATAAATAAATAAATGCAATTTATGTAATATTTTCAACACTC 1928
Db 1447 ANTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1506
Qy 1929 AACATTCATGTTAGGTACTATAAAT 1954
Db 1507 AATAAANAATATTTTATNTNANAT 1532
```

```
RESULT 6
CG753083/c
LOCUS
DEFINITION
P048-1-C01.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
CG753083
CG753083.1 GI:37977199
GSS.
Pristionchus pacificus
Pristionchus pacificus
Bukayota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1896)
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
22835951
12884007
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
source
1..1896
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."
```

```
Query Match 4.6%; Score 93.6; DB 9; Length 1896;
Best Local Similarity 44.1%; Pred. No. 5.1e-07;
Matches 520; Conservative 0; Mismatches 645; Indels 15; Gaps 4;
Qy 531 ACTATACCTTTTGGCCGAATGCTTCTCAAAATGTTTTTATATATGTAATAATATGCCCATCC 590
Db 1833 AATNTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1774
Qy 591 AAGGATAAGTAAATTTCCCGTTTAAACAGCTTTGTTAATAATATATATGTTTACACTTACAAG 650
Db 1773 AATAAAAAATTAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1714
Qy 651 GGATATTCGTAATACITTTTAGACGACAGAGACTAGGTCAAAAATCGACGCTGTTAAAC 710
Db 1713 TATTAATAATNTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1654
Qy 711 AGCCTAGACTTGGTCACTGATATAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAA 770
Db 1653 ATAAATAATTAATNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAT 1594
Qy 771 TGACATTTAAATTTAGAGCTATTAATTAAGTTACTATAATAAAGAGAGAGGTTAGTAAACAG 830
Db 1593 TTAATAATATNTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1534
Qy 831 AAGCAGGTAAAAACAAGAGCTTGGCTGCTGTTTGGTTAGTTGTTGTCAGCTCATCTCTTT 890
Db 1533 AATATTTTAAATATTTTATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1474
Qy 891 AAAAGTAATCTAAACTGATCTAAAG-CACATAGAAATTTAGTACAGGTAAACCTTTTAC 949
Db 1473 TTTTNTTNTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1414
Qy 950 AAGAATTTATATTAACGAAAAATCATTTTATAACATGCTCTCGGCTGTCATTATAATAG 1009
```

Db 1413 TAAATTTTATTTTAAATATTTAAATATATTTAAATTTNANNAATATAAATATATAT 1354
Qy 1010 GGATCACTTACTGATCATCCATTAACCTTGTGTTAAACAAATTCATGAGATATAATAT 1069
Db 1353 AAAAAATATNATAATTAATAATTAATAATATTTTAAATTTTAAATTTTAAATTTAT 1294
Qy 1070 CTTACAAATGAAAGAGGACAAATGCTCTTTGAAAAAACAATAGGTACTCCCTCGTCC 1129
Db 1293 AAAAAATAAAAAAGAAAAATTTATATAAATAATATAATATATNAAAAAATATAAAAA 1234
Qy 1130 CTCTGAAATGATACATATGGAATGGACACGGAGACTAAGAAAAATGTTATAAGTAAATG 1189
Db 1233 AAAAAATATATAATTAATAATATATAAATAAATAAATTTTAAATTAATAATTT 1174
Qy 1190 AGAGTAAAAAGAAAGAAAGAAAGTGGGTAAAGTAGGGGACCCACCAATATATAAT 1249
Db 1173 ATAAATATATATAATATAAATAAATAAATAAATAA-----AAAAAATAAATAT 1123
Qy 1250 GATAGATTTAGAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAATAAT 1309
Db 1122 TATAAATTTAAAAATAAAAATNTATTTTAAATTTAAAAATATAATATATTTATTTAT 1063
Qy 1310 TTACTATTTTGAGAAAGTTTGAATGCTATAGAAATTCAGTGGGACATCCATAAAAAGSAA 1369
Db 1062 TTATATAATTAATTAATATTTTATATATATATATAATAAATAAATTTATAATAAT 1003
Qy 1370 GTGTATAGAAATTAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTGA 1429
Db 1002 TTTTAAATAAATAAATAANATATTTTATTTATTAATTTTATTAATAATATATAAATAA 943
Qy 1430 TTTTGTATTCATAGATTATAATCTA--TGTTATATGATATATAATTTTAAAAATAA 1487
Db 942 AATTTTATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 883
Qy 1488 TACTATATT---AATCTGATAGTCGATTACCGCCTTTTATATATTTTACAATCTAGCT 1544
Db 882 AATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 823
Qy 1545 AATATGAAATTAATCAGTTATCTGAAAGCAATAATATCTTTGTAAGACAGCTCGGTC 1604
Db 822 ATNATNAAAAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 763
Qy 1605 AATGGGAAGTTCATGTGATTCAATAGTTTAAATATAAAGTAATTTTAAATTAATTC 1664
Db 762 ATTATATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 703
Qy 1665 TTATTTTGTTCAGAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1704
Db 702 ATATAAATAATATATTTTATTTTATTTATATATTTATTAAT 663

RESULT 7
CC262481/c
LOCUS CC262481 1202 bp DNA linear GSS 13-MAY-2003
DEFINITION CH261-167M9_Sp6.1 CH261 Gallus genomic clone CH261-167M9,
genomic survey sequence.
ACCESSION CC262481
VERSION CC262481.1 GI:30607397
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1202)
Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine

Email: submissions@wustl.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 30
High quality sequence stop: 105.
Location/Qualifiers
source
1. .1202
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-167M9"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 4.5%; Score 93; DB 8; Length 1202;
Best Local Similarity 47.3%; Pred. No. 6.6e-07;
Matches 393; Conservative 0; Mismatches 420; Indels 17; Gaps 4;
Qy 886 TCCTTTAAAGTAGTAATCTAAAGCATGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTT 945
Db 1174 TATNANAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1115
Qy 946 TTACAAGAAATTTATATTAACGAAAAATCATTTTAACAATGCTCTCGCGCTGTCATATA 1005
Db 1114 ATAAAAATTTTATATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1055
Qy 1006 ATAGGATCATTCTGATCATCTTAACAACTTGTAAACAAATTCAAATGAGATATA 1065
Db 1054 AAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1003
Qy 1066 ATATCTTCAATGAAAGAGGACAAATGCTCTTTGAAAAAACAATAGTACTCCCTCC 1125
Db 1002 ATATNTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 943
Qy 1126 GTCCCTCTGAAATGTATACATAATGGAATGGACACGAGACTAAGAAAAATGATATAAGTA 1185
Db 942 ATATTTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 883
Qy 1186 ATGTAGGTAAAGAAAGAGAAAGAGTGGGTAAAGTAGCGGACCCACCAATATAT 1245
Db 882 ATTAATAATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 823
Qy 1246 AATTGATAGATTTAGAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATATAA 1305
Db 822 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 763
Qy 1306 AAATTTTACTATTTTGAGAAAGTTTGAATGTATAGAATTTGAGTGGGACATCCATAAAG 1365
Db 762 TAATAAATTTAAATATAATATATATATAATAATAATAATAATAATAATAATAATAAT 703
Qy 1366 GAAAGTGTAGAAATTAATGGACAGAGGAGTAATACCTTTATGATATATAAATTTT 1425
Db 702 AAAAA--TAAAAATNTAATTTNAATAATTTTATAATAATAATAATAATAATAATAAAT 645
Qy 1426 GTATTTTTCATTTTCATAGATTAATACTATGTTATATAATATATAATAATAATAAAT 1485
Db 644 TATTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 595
Qy 1486 AATACATATATTAATCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACAGTA 1545
Db 584 AAAAAAATTTTATTTTATATATAATAATAATAATAATAATAATAATAATAATAATA 531
Qy 1546 ATATGAATTAATCAGTTATCTGAAAGCAATAAATATCTTTGTAAACAGCGTTCGGTCA 1605
Db 530 ATATATAATAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 471

```

QY 1606 AATGGGAAGTTCATGTATTCATAGT-TTTAATATAAAGTAATTTTAAATTAATTG 1664
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
470 AAAAAATAATATATATATTTTAAATTTTAAATTTTAAATTTAAATTTAAATTAATTT 411
QY 1665 TTATTTTGTTCAGAAATTTAAATAAATTTATGACATGGGAGTTCA 1714
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
410 TTTTATATAAATAATTTAATATATATATAATAAATAATTTATTTATAAATAA 361

RESULT 8
CNS0021J 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL061936
VERSION AL061936.1 GI:4940214
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

FEATURES
source Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR05N11"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 4.5%; Score 92.4; DB 9; Length 1101;
Best Local Similarity 39.4%; Pred. No. 8.5e-07;
Matches 349; Conservative 65; Mismatches 469; Indels 2; Gaps 2;

QY 868 AGTTGTTGTGAGTCATTCTTTTAAAGTAGTAATGTAATCTTAAGACACATAGAAATT 927
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 AGTGTTCATGAAGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 197

QY 928 TAGTACAGGTTAAACTTTTACAGAAATTTTATTAACGAATATCATTTTATAACATGT 987
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 NAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 257

QY 988 CTCTCGGCTGTCTATTATAATAGGATCACTTACTGTATCTCATTTAAACCTTGTATAAA 1047
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
258 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 317

QY 1048 CAAATTCATGAGATAAAATATCTTACAATGAAAGAGGACATGTCTCTTTGAAAAA 1107
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 377

```

```

QY 1108 CRAATAGGTACTCCCTCCGTCCTCGAATGTATACATATGGAATGGACCGAGACTA 1167
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
378 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 437

QY 1168 AGAAAAATGTATAAGTAATGTAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1227
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
438 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 497

QY 1228 CGGGACCCCAATATATTAATTCATAGATTTCAGAAAAGTAGTTGAAAAGTAGTCGGT 1287
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 557

QY 1288 GGGATTTTATATATATAAATAATCTATTTTGGAGAAAGTTTGGAAATGTATGAATGA 1347
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
558 TWTTAATTTTAAATTTTAAAWAAATTTAATAAATAAATAAATAAATAAATAAATAA 617

QY 1348 GTGGGACATCCATAAAGGAAAGTGTATAGAAATTAATGGACAGAGGAGTAGTAATCCTT 1407
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
618 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 677

QY 1408 TATGATATAT-AAATTTTGTATTTCATAAGATTATAAATCTATGTTATAATG 1466
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
678 TWTATWAAATWAAATTTTAAAWAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 737

QY 1467 ATAATATAATTTTAAATAATACTATATTAATCTGATTAGTCGATACCGCTTTTAT 1526
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
738 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 797

QY 1527 AATTTTACAAATCTGAGTATATGAATAATCAGTATCTGAAAAGCAATATATCTTT 1586
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
798 TTTAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 857

QY 1587 GTAACACAGCGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1646
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
858 TTAATTTTAAATTTTAAATAATACTATTAATAATTTTAAATAATACTATTAATAAT 916

QY 1647 TAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1706
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
917 AAATGKGAATGKKKGKTKARKKKRGKTTAARATDGTAAAGAKRATAAKTKTKATKK 976

QY 1707 GAAGTTCACGGGATCATTTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1751
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
977 TAAAAAADKAKGAAAAATAAWGAARGKATGAAAAAGTGKDAATAT 1021

```

```

RESULT 9
CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of

```


[illegible]

	Qy	799	GTTACTATAAATAAAGAGGGTGTAGTAACAACAGAAGCAGGCTAAAAACAAGAGCTTGCTGC	858
	Dd	221	ATTAAAAATTAATATAATTTAAAATTTAAATTTATTAATAATATAATAATATAATAACAATA	280
	Qy	859	TGTGTGTTTAGTTGTTCGTGAGCTCATTTCTTTAAAAGTAATGTAAAACTGATCTAAAGCAC	918
	Dd	281	ATAAAATTATTAATTAATAATAAAATATATACATATTAAATAATTTCAAAAATTTATAAAATTA	340
	Qy	919	ATAGAAATTTAGTACAGGTTTTAAAACTTTTACAAGAAATTTATATATAAACCGAAAAATCATTTT	978
	Dd	341	TATAAATTAATAATTTATTTAAATATATTAATAATAATTTTATTTATATAATAATAAAATATA	400
	Qy	979	ATAACATGCTCTCGGGCTGTCATTTATAATAGGAATCACCTTCTGATCATCCATTAAAAACC	1038
	Dd	401	ATTATATAATAAAAAATATAATATAATAATAATTTTATTTAAATAAAAAAATTTAAATAATA	460
	Qy	1039	TTTGTTAAAACAAATCCAATGAGATAAAATATCTTACAATGAAGAAGGACAATGCTCT	1098
	Dd	461	TTTATTATATTTTATTTATAAAATTTATAAAATTTATTAAATTAATTTTAAATAATAATAA	520
	Qy	1099	TTGAAAAAACAATAGTACTCCCTCGCTCCCTCGAATGTATACATATGGATGGACA	1158
	Dd	521	ATATTAAACTATATATTTTAAATAATTTAAATAATATATTAAATAATAATAATATTATATA	580
	Qy	1159	CGBA - GACTAAGAAAAATGTATAAAGTAATGTAGAGTAATAAAGAGAAAGAGAAAGTG	1217
	Dd	581	TAAATTTATTAAAAATAAAAAATTTAAAAATAATAATTATAAAAAATAATAAAAAATAT	640
	Qy	1218	GGTAAGTAGCGGGACCCACCAATATATATTCATAGATTTTACGAAAGTAGTTGAAAGTA	1277
	Dd	641	ATTAATAATAAAATTAATATTTAAATAATTAATTAATAAAATAAAATAAAATTTATATAA	700
	Qy	1278	GTGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTTCGAGAAAGTTTTCAAATGT	1337
	Dd	701	ATTAATTTTAAATTTTATATAATAATTAATTTATAATAATTTTAAAAATTTTATAATTAAT	760
	Qy	1338	ATGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAATTAATGGGACAGAGGGA	1397
	Dd	761	AATATATATAATTTTAAATAAATTTAAATATAATTAATAATAATTTATAATTTAAAAATAA	820
	Qy	1398	GTAATACCTTTATGATATATAAATTTTGGTTATTTTGGATTTTCATAGATTTATAAACTAT	1457
	Dd	821	AAAATATATAAAATTTATATAAATTTTATTTATATAAAAAAATAATAAAAAATAATAAT	880
	Qy	1458	GTTATAATGATAATATAATTTTAAAAATAACTATATATAATTAATTTCTGATTAGTCGATTACC	1517
	Dd	881	TATAAAATTAACATAAATAATATAATTTTAAATTTATTAATTTAAATTTAAAAATAATAA	940
	Qy	1518	GCCTTTTATAATTTTACAATCTAGTGTAATATGAATAAATCAGTTATCTCGAAAGCAAAT	1577
	Dd	941	ATATTTTATA - --TAATAAAAATATATTAATTTTATAAAAAATAAAAAATAATAAAAAAT	997
	Qy	1578	AAATCTTTGTAAAAACGCGTTCGGTCAAAATGGGAGGTTCATGTGTATTCCAATAGTTTAA	1637
	Dd	998	AAATTTTATATAAAAAAATAAAT - -ATAAAATTTATAATTTTATAAATTTCTATTTAATA	1055
	Qy	1638	ATATAAAGTAAATTTTAAAAATTAATTCGTATTTTGGTTTCAGAAATTTAAAAATAATTA	1697
	Dd	1056	ATAATATATTAAAAAATTTAATTTATTATATTAATTAATAATAATAATAATAATAATAA	1115
	Qy	1698	T	1698
	Dd	1116	T	1116

RESULT 15	
AG346840/c	
LOCUS	AG346840
DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-142C12.T7, genomic survey
	1696 bp DNA linear
ACCESSION	AG346840
VERSION	AG346840.1
KEYWORDS	GI:47920150
	GSS.

```

SOURCE
ORGANISM
Mus musculus molossinus
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Hattori M., Toyoda, A., Noguchi H., Kojima, T. and Sakaki, Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1696)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9114, Fax: 81-45-503-9170)
COMMENT
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
source
Location/Qualifiers
1..1696
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-142C12.T7"
/sex="male"
/tissue types="mixture of kidney and spleen"
/clone lib="MSMg01 Mouse Male BAC Library"

```

[illegible]

```
Qy 1135 AATGTATACATGATTCGACACGGAGACTAAGAAATGTATTAAGTAATCTAGAGT 1194
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 879 TAATAAAATAAATAATATAATAAACA-TATAAAAAATATATAAAAAATATACAAA 821
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1195 AAAAAGAAAGAGAAAGTGGTAAAGTAGCGGACCCACCAATATATAATTGATAG 1254
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 820 ATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 761
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1255 ATTTAGAAAGTAGTTGAAGTAGTGGTGGTGGGATTTTATATATATAAATTTACT 1314
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 760 AAATA-AACATAACAAATAATATAAAAAAATAAAAAACATATAAATAATATAAAAAATA 702
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1315 ATTTGAGAAAGTTTGAATGTATAGATTTGAGTGGGACATCCATAAAGGAAAGTGA 1374
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 701 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 642
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1375 TAGAATTAATGGACAGAGGGAGTAACTCTTTATGATATATAAATTTTGTATTG 1434
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 641 TATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 582
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1435 ATTTCAATAGATTATAAATCTATGTTTAAATGATATATAATTTT-----AAAAATAATAC 1490
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 581 AAAAAACAA-ATTAAAAATATATAAATTAATAATAAATCTTTAAAAAATAAATAAATA 523
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1491 TATATTAATTTCTGATTTAGTCGATTCACGCTTTTATAATTTTACAACTACTGAGTAATG 1550
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 522 AATTATAATAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 463
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1551 AATAAATCAGTTATCTGAAAGCAATAATAATATCTTTGTAACAACAGCGTTCGGTCAAAATGG 1610
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 462 AAGAAATTAATAAATAATACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 403
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1611 GAAGTTCATGTTTCAATAGTTTAAATAAAGTAAATTTTAAATTAATTTGTTATTT 1670
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 402 AATAAATAATCAACTAAAAACAAAAAACTAAAAAATAAATAAATAAATAAATAAATAAATAAATA 343
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1671 TTGTTTCAGAAATTTAAATA 1691
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 342 TTATCTCAAAAAATAAATAAATA 322
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 14, 2005, 20:23:05
Job time : 6856 secs

THIS PAGE BLANK (USPTO)